4. TAG 2 (IS-95-BASED CDMA SYSTEM) TESTING

This section describes the test plan, methodology, and results for the technology field trial conducted by TAG 2. The technology tested was the PCS variant of the IS-95 cellular CDMA standard. For this field trial, rate set 1 and the IS-96A codec were used. The three-chip base station technology was used. All TAG 2 equipment was provided by Qualcomm, Incorporated. The TAG 2 field tests examined area coverage, handoff, and voice quality of the system.

The information presented in this section is taken from [2]. The reader is referred to [2] for a more complete and detailed presentation of the TAG 2 technology field testing at the BITB.

4.1 TAG 2 Test System Configuration

The block diagram of the test system configuration is shown in Figure 4.1. The test system consisted of the three cell sites (WCO, TMCO, and GMM) and one Qualcomm Telecommunications Switching Office (QTSO) located in Boulder, Colorado. The system functioned as a closed system as calls to and from the public switched telephone network (PSTN) were not allowed during the trial. The QTSO was located at the BITB 28th Street laboratory. As shown in Figure 4.1, the T1 line from each of the TMCO and WCO base stations was connected to a digital cross-connect and fiber multiplexer to provide an OC3 optical link between the base stations and the QTSO. As in the TAG 5 test system configuration, an LOS microwave link was used to provide a partial T1 link between the GMM base station and the WCO base station. For all of the TAG 2 testing, the uplink (mobile transmit) center frequency was set to 1872.99 MHz and the downlink (base station transmit) center frequency was set to 1952.99 MHz.

4.1.1 Simulated Interference

All data were collected in the presence of simulated interference. The interference was simulated on the uplink by the other user noise simulator (OUNS) and on the downlink by the orthogonal channel noise simulator (OCNS).

Uplink capacity was simulated based on the effects of interference from users in the same cell and users in surrounding cells. The noise generated by these other users was modeled as a Gaussian random process simulated by a single "big fictitious" user. The resulting simulated noise of N simultaneous users was injected at the inputs of the IF amplifier. The N simulated users were assigned a desired E_b/N_o (energy per bit to noise density ratio) set point to be maintained (8.5 dB in this testing). This set point was normally set higher than the QTSO recorded E_b/N_o set point as a safety margin to ensure that the interference was at least that of the desired users.

The downlink capacity was simulated by modulating a random information stream on unused Walsh codes. The number of Walsh codes used was fewer than the number of simulated users. To obtain the equivalent power of the simulated number of users, the power in each Walsh code

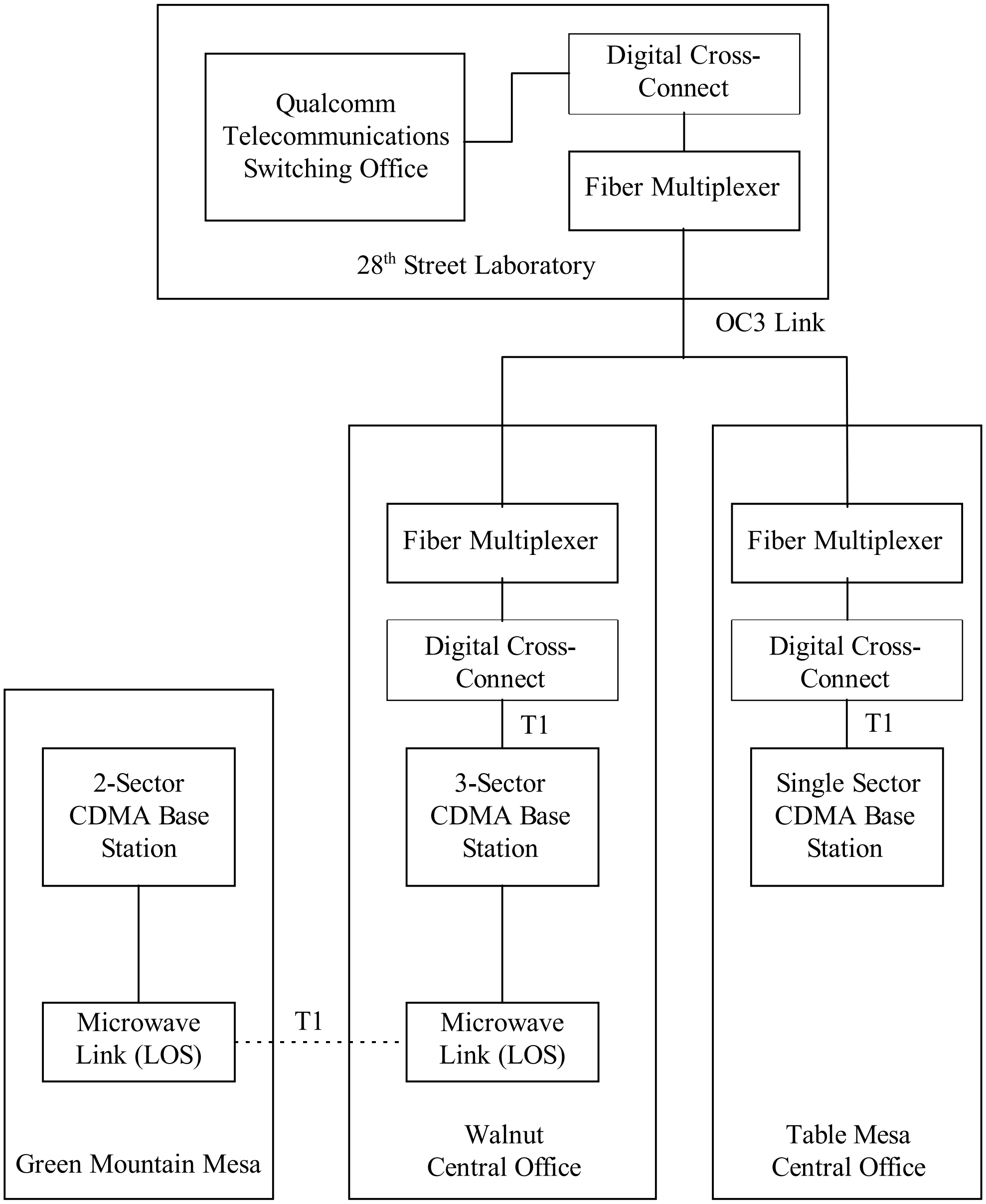

Figure 4.1. Block diagram of the TAG 2 test system configuration.

used for simulation was increased so that the total power was identical to that if a Walsh code was used for each user.

4.1.2 Markov Calls

During the testing of the TAG 2 technology, "Markov" test calls were used to collect link performance statistics. During a Markov test call, the mobile unit and the vocoder/selector (within the QTSO) use synchronized pseudorandom generators and statistical models to generate and verify the data rate and data bits over the air. The data rate of a Markov test call can be set to fixed or variable rates. For a fixed rate, the data rate is set to one of the following values for the duration of the call: full rate (9,600 bps), half rate (4,800 bps), quarter rate (2,400 bps), or eighth rate (1,200 bps). A variable rate Markov call uses a combination of the above rates according to the statistics of a typical voice conversation. The variable rate Markov call produces a voice activity of 40%.

4.2 Calibration

The calibration performed at the field trial for TAG 2 was not a formal calibration of the equipment but was instead a verification of the accuracy of the data measurement reports. Formal calibration of this equipment was carried out by the manufacturer at the factory.

The mobile unit, used for all mobile data logging, was checked for accuracy of the RSS values reported. The reported RSS is the total received signal power within the IF bandwidth and is detected prior to correlation (de-spreading). The base station was connected to the mobile unit via a coaxial cable for this calibration. A known signal level was injected into the mobile unit and compared (taking into consideration the transmission loss) to the RSS value reported on the mobile diagnostic monitor (an instrument that provides information about the operation of the mobile unit). The mobile unit RSS, as reported by the mobile diagnostic monitor, was always within 3 dB of the actual RSS.

The mobile diagnostic monitor gives only a coarse report of the RSS. A more accurate calibration curve for the mobile unit is stored in the data analysis tool (proprietary software from the manufacturer) used for final data processing. The independent observer was provided with a copy of the mobile station factory calibration curve.

A qualitative examination of the base station transmitter characteristics was also performed; however, the results of this examination are not presented here. The base station RSS, and other parameters such as the base station received E_b/N_o, were not examined during this calibration.

4.3 Area Coverage Testing

Measurements to show area coverage were taken with the mobile unit[4] located in a mini-van as in TAG 5 testing. The mobile unit was mounted inside the van on the same wooden structure as used in the TAG 5 testing. The measurements were taken by driving along routes (radials) away from the cell site and filling in-between the radials as time permitted. Markov calls with a 40% voice activity factor were used to generate link performance statistics. The OCNS and the OUNS were set to simulate 10 active users in an embedded cell (i.e., a cell surrounded by other operational cells).

[4] In the JTC test report on TAG 2 testing, a distinction was made between portable and mobile handsets. The mobile handset used an antenna mounted outside of the van while the portable handset used an antenna located in the measurement van. In this section of this NTIA Report, the term mobile unit implies a handset with an antenna located in the measurement van. For purposes of brevity and because it was not a formal part of the JTC PCS technology field trials, TAG 2 measurements performed with an antenna located out of the measurement van are not discussed in this NTIA Report.

Only one cell site was activated at a time during area coverage testing. During the area coverage testing for each cell, all other cell sites were shut down. Softer handoff was allowed between the sectors of the active cell. Softer handoff occurs when the mobile unit initiates communication with a new active serving sector of the same cell while maintaining communication with the currently active serving sector. For each drive route, a Markov call was initiated close to the cell site. Collection of mobile data and cell site data was initiated as the measurement van traveled away from the cell site along the drive route. (In some cases, shadowing or poor coverage caused the call to be dropped on the route; in those cases, the call was reinitiated if the pilot carrier signal[5]-to-interference ratio (E_c/I_o) improved in a short distance.) At the end of the route, the data collection was stopped and the data were saved to disk. The data were collected at the mobile unit and at the QTSO. The data from the different sources were then collated with respect to time and combined to provide a file that included GPS location, velocity, and time; downlink RSS; mobile transmit power; uplink and downlink frame error rate (FER); QTSO target E_b/N_o; uplink E_b/N_o; received pilot E_c/I_o at the mobile unit; and other system parameters. All data were recorded once each second except for the FER statistics. The FER statistics were calculated on a 500-frame basis that corresponds to a minimum 10-s sample time.

In addition to driving from the cell site outward, some routes were also driven in the opposite direction, i.e., driving towards the cell site. These routes were excluded from the analysis in this section for consistency. However, the outbound routes show a lower RSS (by approximately 6 dB) than the inbound routes due to shadowing effects from passengers inside the vehicle and the tinting of all windows in the measurement van except the windshield.

4.3.1 TMCO Area Coverage Data

The data were collected as explained in Section 4.3. Only the north-facing sector was active for this cell site.

Downlink RSS as a function of distance is shown in Figure 4.2. This figure includes the overall data for this cell, excluding RSS values less than -104 dBm. Those values were excluded because -104 dBm is the threshold level for normal system operation. The large variation in RSS is partially due to the fact that the data were taken outside of the north sector coverage area. On the other hand, relatively strong signals (approximately -76 to -82 dBm) exist far away from the cell (approximately 13.5 mi). Those signals were recorded in an area where LOS propagation existed between the base station and mobile unit.

A rough estimate of the coverage area was determined by assuming that an RSS of -100 dBm or greater is desired. The measured RSS data along all of the routes driven within the cell were used to determine the coverage area. Due to the irregularity of the terrain, the RSS varied significantly along the TMCO routes, crossing the -100-dBm level several times before finally staying below -100 dBm. The point along each route where the RSS first dropped below -100 dBm was used to define the coverage boundaries. For this case, the coverage boundaries were approximately 5.7 mi due north-northeast, 3.75 mi due northeast, and 4.44 mi due east. Since the TMCO cell is at a higher elevation than the WCO cell, the TMCO cell had greater coverage in the northern-northeastern direction than the WCO cell.

---

[5] The pilot carrier signal is an unmodulated spread spectrum signal that is transmitted from a CDMA base station.

Figure 4.3 shows the histogram of downlink RSS values. Again, RSS values less than -104 dBm were excluded. The mean RSS is -91.8 dBm with a standard deviation of 10.0 dB.

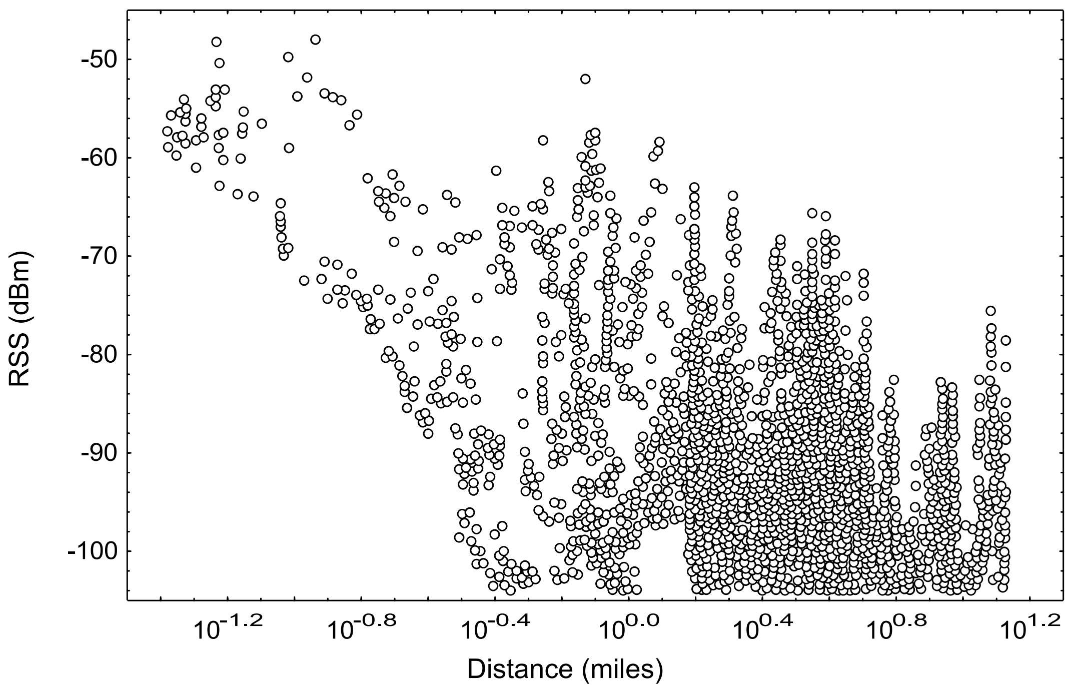

Figure 4.2. Downlink received signal strength (RSS) vs. distance (TAG 2, TMCO cell).

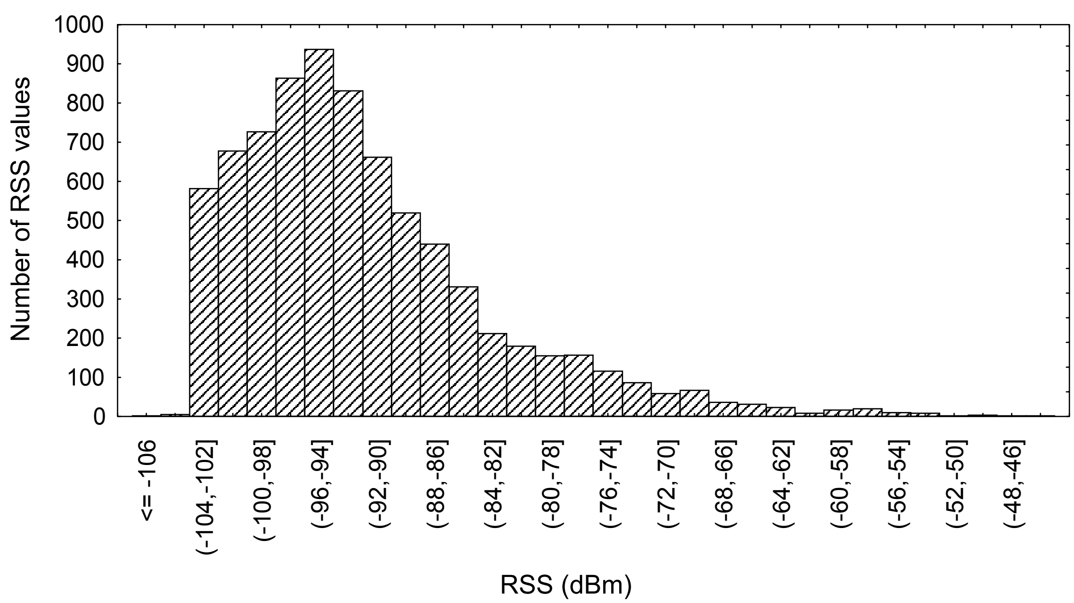

Figure 4.3. Histogram of downlink received signal strength (RSS; TAG 2, TMCO cell).

Figure 4.4 shows the histogram of uplink transmit power. The maximum transmitted power is +26 dBm. Approximately 1.6% of the data points had a transmit power greater than +25 dBm. That indicates that 1.6% of the time, the mobile unit was transmitting at or near its full power capacity. The mobile unit's transmit power increases when its corresponding RSS decreases. The RSS decreases as the mobile unit moves farther away from the cell site or if there is signal loss due to shadowing. The areas where the mobile unit was transmitting at its full power capacity were generally the areas close to the edge of coverage for this cell.

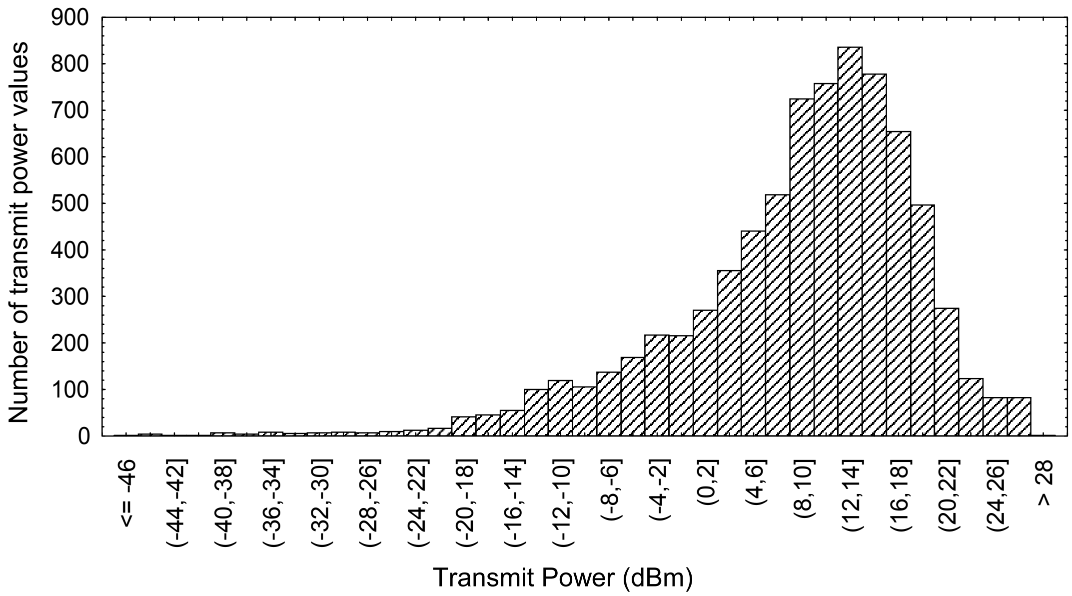

Figure 4.4. Histogram of uplink transmit power (TAG 2, TMCO cell).

The QTSO target E_b/N_o was set by the open loop power control to provide an uplink FER of 1%. When the mobile unit moves outside of the coverage area and can transmit no more power than its maximum rating, the QTSO target E_b/N_o reaches its maximum value (set to 8 dB for this test). When the mobile unit cannot transmit enough power to meet the target E_b/N_o the uplink FER increases above the 1% target. This is clearly seen in Figure 4.5 (a plot of uplink FER vs. QTSO target E_b/N_o). The target E_b/N_o varies between its minimum and maximum values (set to 5 and 8 dB, respectively in this test) in order to keep FER at or below 1%. As the mobile unit moves outside of the region of good coverage, the QTSO target E_b/N_o reaches its maximum and the uplink FER increases until the call is dropped. (Note that the actual uplink E_b/N_o decreases.)

Figures 4.6 and 4.7 show the histograms of the uplink and downlink FER, respectively. These histograms show that most of the FER's for both the uplink and downlink were less than or equal to 1%.

One way to estimate the link balance is to compare the uplink FER and downlink FER. However, the uplink FER and downlink FER were sampled at different times. The time difference varied from one to several seconds. The measurements at the mobile unit

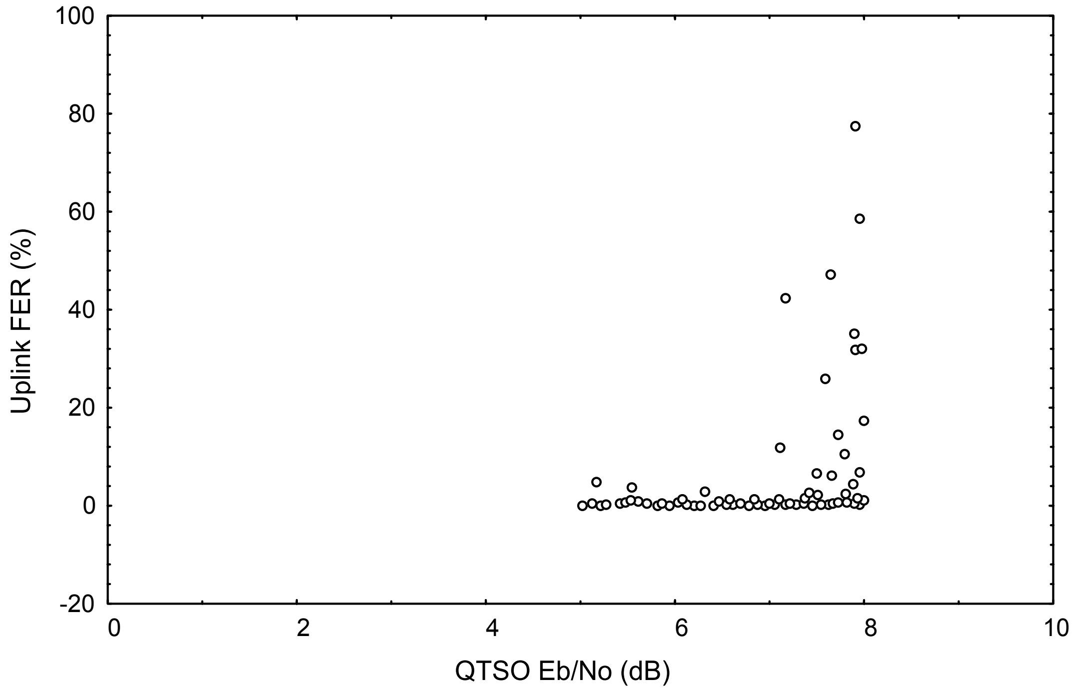

Figure 4.5. Uplink frame error rate (FER) vs. Qualcomm Telecommunications Switching Office (QTSO) target E_b/N_o (TAG 2, TMCO cell).

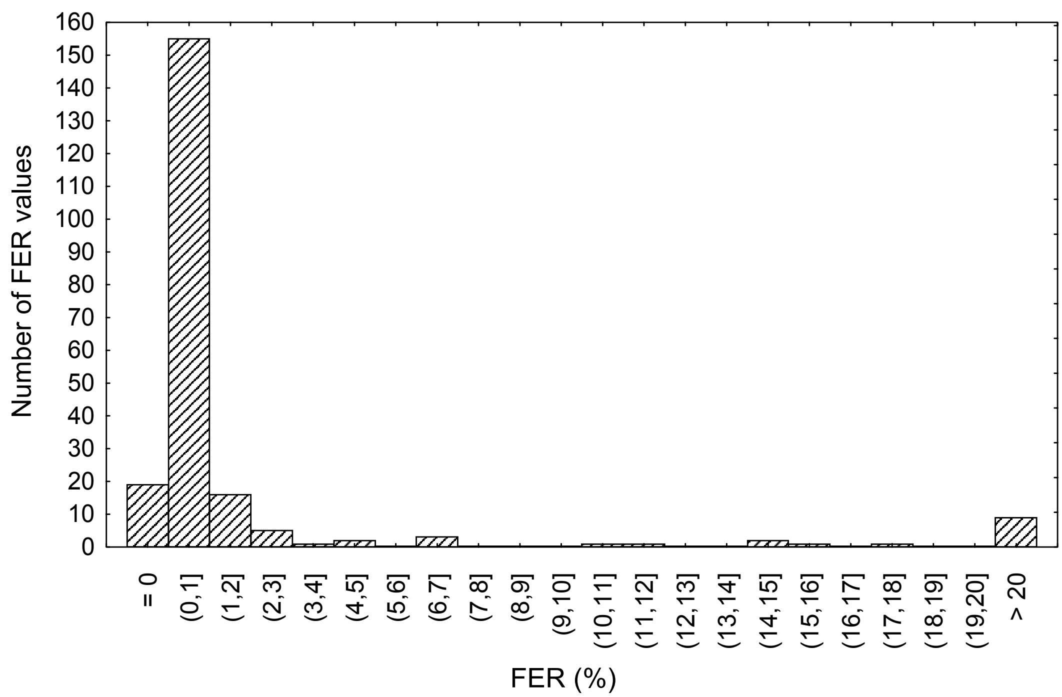

Figure 4.6. Histogram of uplink frame error rate (FER; TAG 2, TMCO cell).

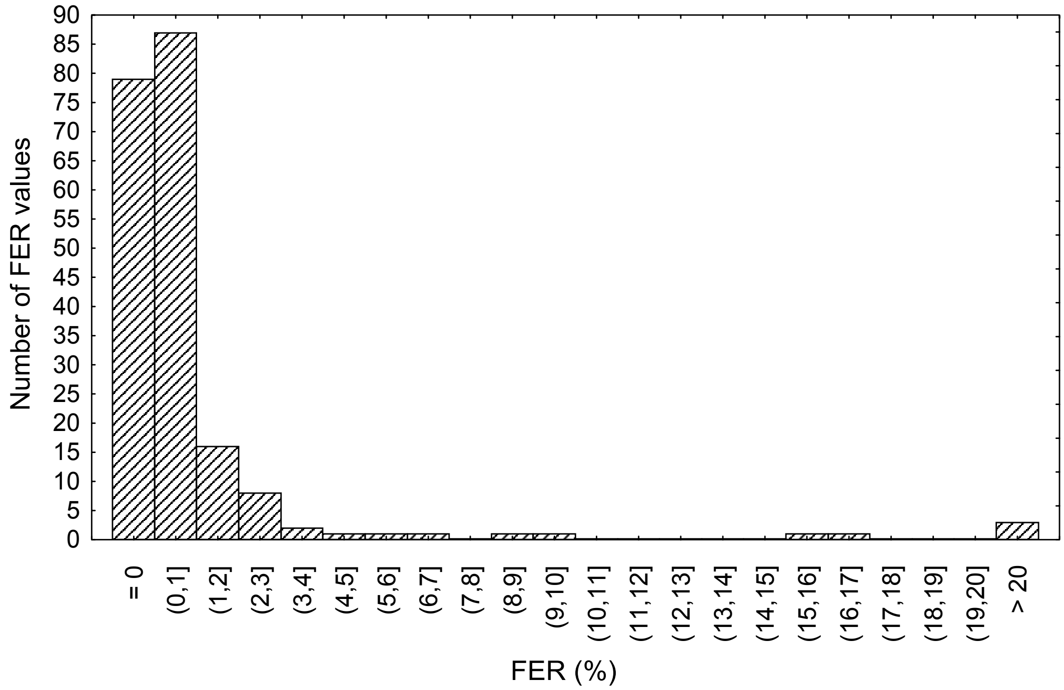

Figure 4.7. Histogram of downlink frame error rate (FER; TAG 2, TMCO cell).

correspond to speeds varying from a full stop (at the traffic lights) to 55 mph. Due to the difference in sample times, the uplink and downlink FER measurements were taken at different physical points — up to 400 ft apart. In order to make a fair analysis of the link balance, it is essential to collect uplink and downlink data at the exact same spot (within a few feet). Therefore, while it is possible to analyze the link balance on an averaged basis, the link balance analysis was not performed for the JTC TAG 2 technology field trial.

4.3.2 WCO Area Coverage Data

The data were collected in the WCO cell in the same manner as in the TMCO cell. This procedure is explained in Section 4.3. All three sectors were active for the WCO cell.

Downlink RSS as a function of distance is shown in Figure 4.8. The data included in this figure are the overall data for this cell, excluding RSS values less than -104 dBm. Those values were excluded because -104 dBm is the threshold level for normal system operation. As with the TMCO area coverage, there is a large variation in the RSS for a given distance. This is due to shadowing in the light urban environment and the choice of routes. Note that there exists a single RSS value of approximately -100 dBm 8 mi away from the WCO cell site. This maximum coverage distance for the WCO cell is significantly less than the maximum coverage distance for the TMCO cell (13.5 mi).

A rough estimate of the coverage area was determined by assuming that an RSS of -100 dBm or greater is desired. The measured RSS data along all of the routes driven within the cell were used to determine the coverage area. The point along each route where the RSS first dropped below -100 dBm was used to define the coverage boundaries. For this case, the coverage boundaries were 1.4 mi due north, 4.5 mi due east, and 1.6 mi due southeast and due west. Once the RSS dropped below -100 dBm in these directions, it generally stayed below that value. Due to the lower elevation of the WCO cell (relative to the TMCO cell) and the terrain profile, this cell had a smaller total coverage area than the TMCO cell.

Figure 4.9 shows the histogram of downlink RSS values. Again, RSS values less than -104 dBm were excluded. The mean RSS is -89 dBm with a standard deviation of 13.5 dB. Figure 4.10 shows the histogram of uplink transmit power. The maximum transmitted power is +26 dBm. Approximately 2.8% of the data points had a transmit power greater than +25 dBm. That indicates that 2.8% of the time the mobile unit was transmitting at or near its full power capacity. The areas where the mobile unit was transmitting at its full power capacity were the areas close to the edge of coverage for this cell or where significant shadowing occurred.

Figure 4.11 shows a plot of the uplink FER vs. the QTSO target E_b/N_o. The target E_b/N_o varies between 5 and 8 dB in order to keep FER at or below 1%. As the mobile unit moves outside of the region of good coverage, the QTSO target E_b/N_o reaches its maximum and the uplink FER increases until the call is dropped.

Figures 4.12 and 4.13 show the histograms of the uplink and downlink FER, respectively. As seen for the TMCO cell, these histograms show that most of the FER's for both the uplink and downlink were less than or equal to 1%.

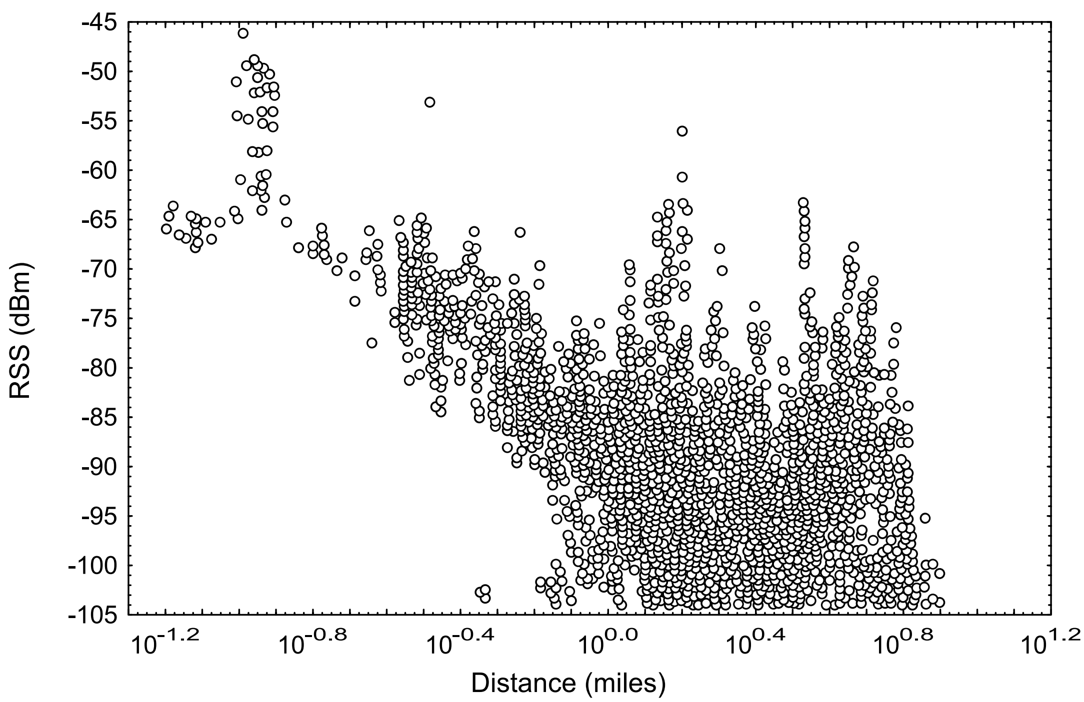

Figure 4.8. Downlink received signal strength (RSS) vs. distance (TAG 2, WCO cell).

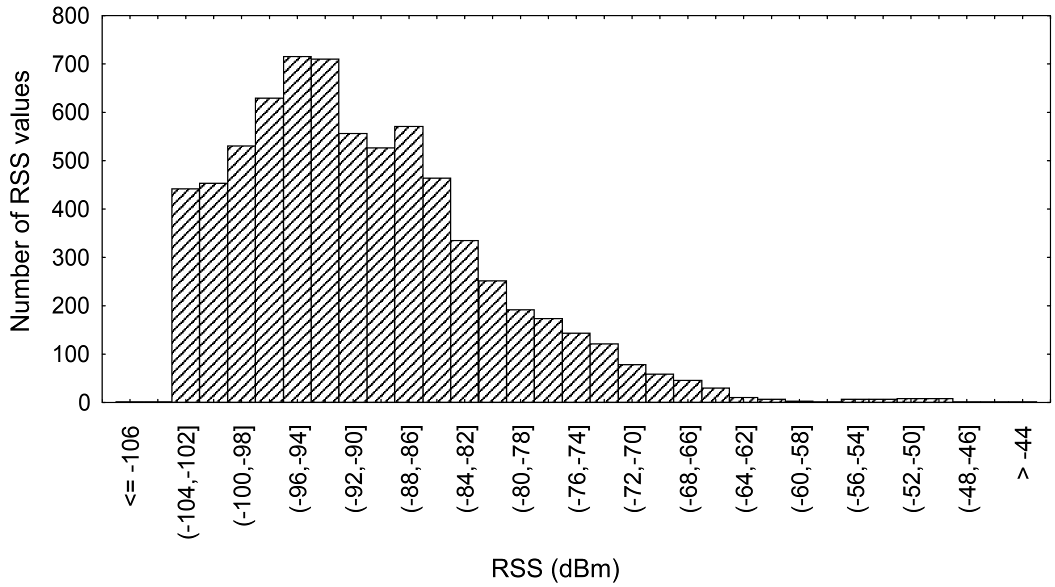

Figure 4.9. Histogram of downlink received signal strength (RSS; TAG 2, WCO cell).

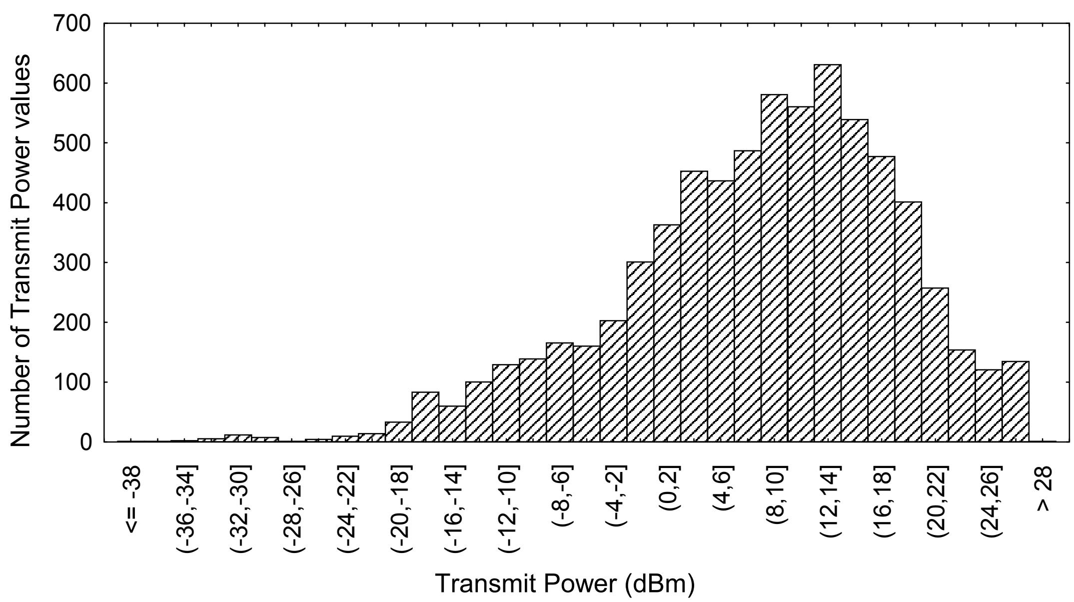

Figure 4.10. Histogram of uplink transmit power (TAG 2, WCO cell).

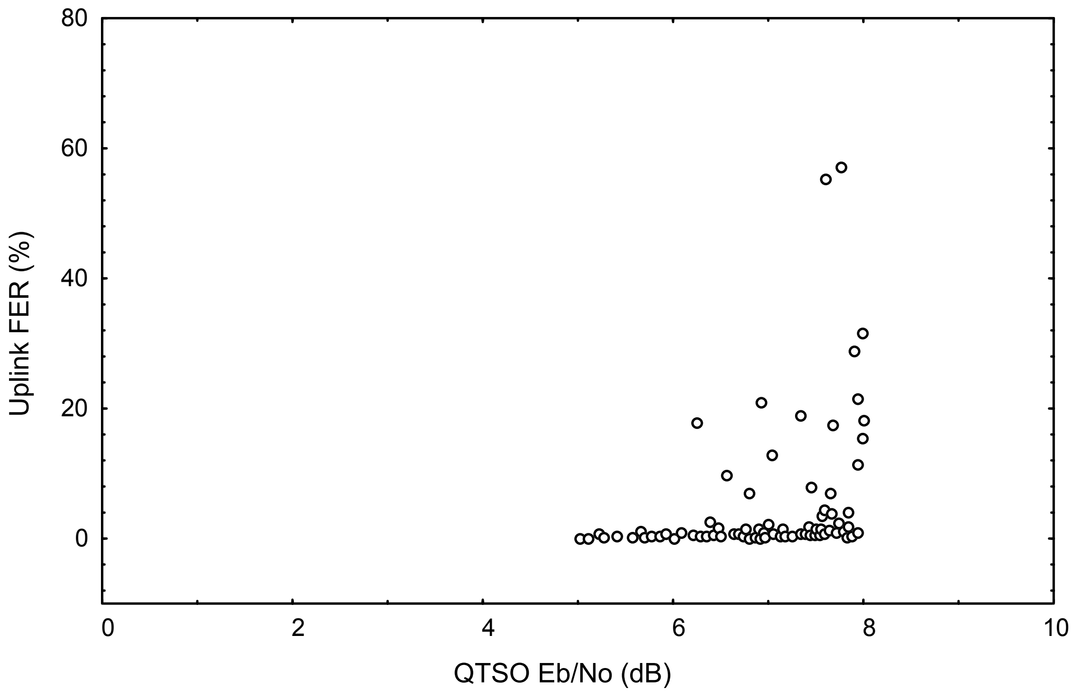

Figure 4.11. Uplink frame error rate (FER) vs. Qualcomm Telecommunications Switching Office (QTSO) target E_b/N_o (TAG 2, WCO cell).

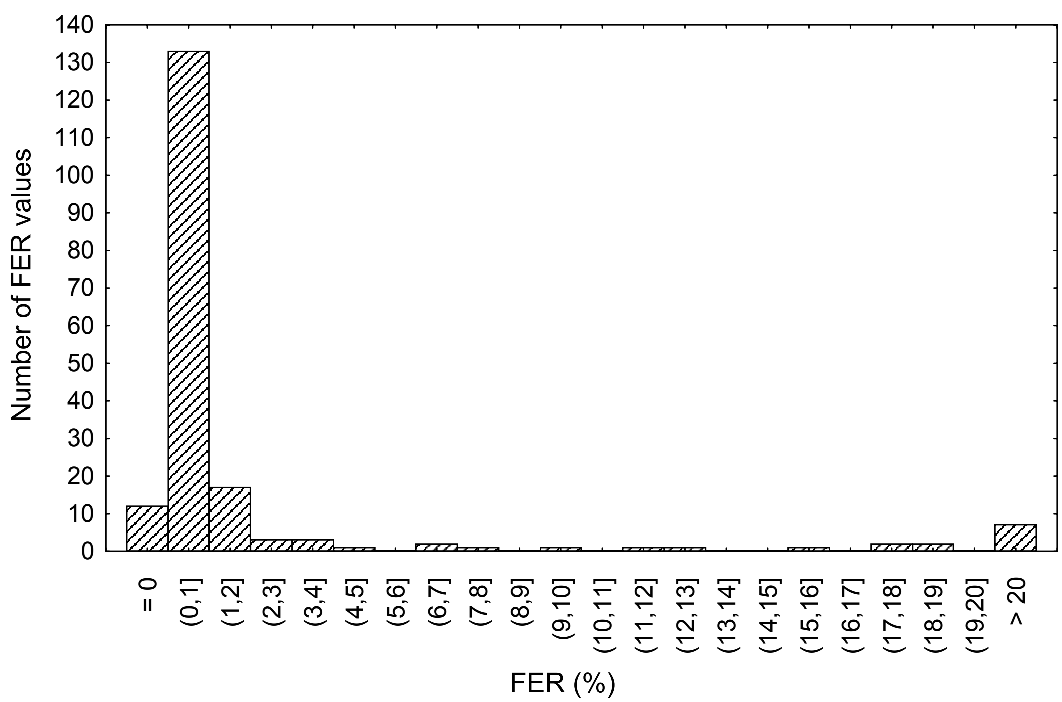

Figure 4.12. Histogram of uplink frame error rate (FER; TAG 2, WCO cell).

4.4 Handoff Testing

The north sector for the TMCO cell site and all three sectors of the WCO cell site were activated for the handoff testing. Soft handoff was employed. Soft handoff[6] occurs when the mobile station initiates communication with a new serving cell or sector while maintaining communication with the currently active serving cell or sector. Handoff testing was performed by driving the measurement van and collecting data along routes that would pass through the coverage area of both the TMCO and WCO cell sites and each active sector within the WCO cell site. Markov calls were used for the handoff testing. The system handoff parameters were set as follows:

Add Pilot Threshold (T_ADD) = -14 dB
Drop Pilot Threshold (T_DROP) = -16 dB
Active Set vs. Candidate Set Threshold (T_COMP) = 5 dB
Drop Pilot Timer Value (T_TDROP) = 6 s.

An explanation of how these parameters affect handoff is quite detailed and can be found in [3]. Data analysis for handoff testing for the TAG 2 technology consisted of determining the handoff state as a function of geographic location and the percentage of time the network was in a particular handoff state (no handoff, soft handoff, softer handoff, etc.).

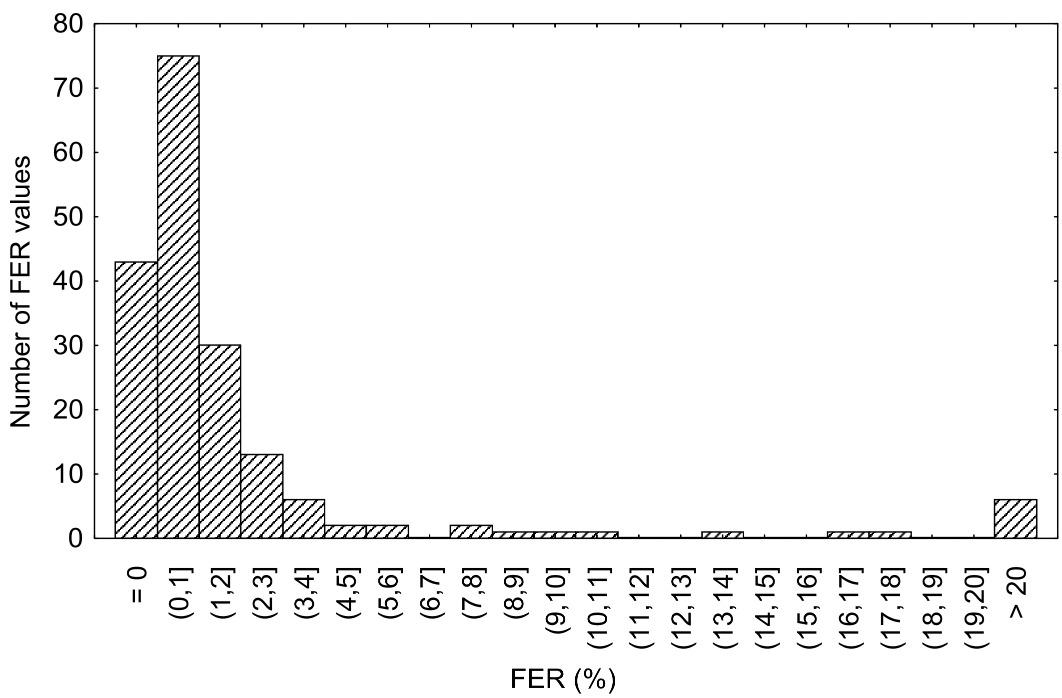

Figure 4.13. Histogram of downlink frame error rate (FER; TAG 2, WCO cell).

[6] More formally, soft handoff occurs between two active cells, softer handoff occurs between two active sectors of the same cell, soft-softer handoff occurs between two active sectors of the same cell and another active cell, and softer-softer handoff occurs between three active sectors of a given cell.

The results of the analysis of handoff state as a function of geographic location showed that no handoff was observed in the immediate vicinity of both the TMCO and WCO cells. Soft handoff was present in the open areas east of Boulder, where both the TMCO and WCO cells had overlapping coverage. Softer handoff was pronounced in areas where the signal from neighboring sectors in the WCO cell were overlapping. Soft-softer handoff occurred in the open areas along Foothills Parkway, east-northeast of Boulder. In this area, two sectors from the WCO cell overlap and there is LOS propagation to the TMCO cell. There are a few other isolated areas where soft-softer handoff occurred, but in a less consistent manner than along Foothills Parkway.

The percentage of time during the handoff testing that the system was in a particular handoff state was then determined. No handoff occurred 37% of the time, soft handoff occurred 40% of the time, softer handoff occurred 9% of the time, and soft-softer handoff occurred 14% of the time.

The percentage of locations in soft handoff is quite high. This occurred because the coverage area for the WCO cell was contained within the coverage area for the TMCO cell (i.e., the WCO cell was an embedded cell in this configuration). Soft handoff occurred frequently because the pilot carrier signals from both the TMCO and WCO cells had a high E_c/I_o. Due to the light load of the system, the pilot carrier signals have a high E_c/I_o. With increasing load, the E_c/I_o of each of the pilot carrier signals decreases. The E_c/I_o of the pilot carrier signal for the original cell sector decreases along with the E_c/I_o of the pilot carrier signals from cell sectors other than the original cell sector. Because the E_c/I_o of the pilot carrier signals from cell sectors other than the original cell sector decreases, there is less potential for handoff to occur and the percentage of measurement locations with no handoff increases. (The manufacturer stated that the system was not optimized for the given test configuration and cell spacing. Instead, the system used the same settings used in other Qualcomm CDMA tests.)

4.5 Interference Testing

Separate interference testing was not performed during the TAG 2 testing because all measurements were made with simulated interference as described in Section 4.1.1.

4.6 Voice Quality

As discussed in Section 3.6, two types of voice quality measurements were made for the PCS JTC technology field trials in general: quasi-stationary measurements and handoff measurements. Both types of measurements were performed for the IS-95-based CDMA (TAG 2) technology.

4.6.1 Quasi-stationary Measurements

Voice recordings and various objective measures including RSS, uplink and downlink FER, and QTSO target E_b/N_o were collected at locations on a 0.5-mi grid that encompassed the expected coverage area for the TMCO and WCO sites. Measurements were taken at 69 of the 82 locations that were identified for the quasi-stationary measurements as discussed in Section 3.6.1. The specific locations used are shown on the map in Figure 4.14. Both the TMCO and WCO cells were activated for these measurements. (In the TAG 5 testing, recall that only one cell was activated at a time.) At each location data were collected as the measurement van traveled at one of two speeds. The vehicle traveled either 10 m or 100 m over the sample time. The particular vehicular speed used at each location (distance traveled over a fixed sample time) is shown on the map in Figure 4.14.

The measurements were taken at each location by establishing a call between the mobile and landline telephones. While the measurement van was in motion, an audio source tape was transmitted over the uplink and downlink simultaneously. The source tape transmitted over each link was the same as that used for TAG 5 testing (see Section 3.6.1).

The received voice transmissions were recorded on digital audio tape simultaneously at the receiver for the uplink and at the receiver for the downlink. The recorded voice segments were then digitized with 16-bit resolution at a 22-Ksample/s rate and stored on a hard disk drive. At each location, the objective measures including downlink RSS; uplink and downlink FER; QTSO target E_b/N_o; and GPS location, velocity, and time were collected during a separate measurement performed after the voice transmissions were recorded. The separate measurements used full-rate Markov calls to approximate the 75% voice activity of the source tape.

For the quasi-stationary measurements, voice quality of the voice segments was determined by both mean opinion score (MOS) and expert listener techniques. The following sections discuss these techniques and present the results based on the application of these techniques.

4.6.2 Mean Opinion Score Assessment

To accomplish the MOS testing, a pool of 31 subjects was recruited from the Boulder, Colorado area. Each of the following age groups were represented within the subject pool: 18-25, 25-35, 35-45, 45-55, and those over 55 years of age. There were an equal number of male and female subjects. The subjects were cordless, noncellular telephone users.

Four groups consisting of seven or eight subjects from the subject pool were formed. The subjects were asked to rate voice segments by answering the three questions listed in Section 3.6.2 after each segment was presented.

First the subjects in each of the four groups were presented 10 practice voice segments[7] to rate. The practice segments included two 64-kbps wireline voice segments and six segments recorded over a speech codec in Qualcomm’s lab. The 64-kbps wireline voice segments consisted of one good quality voice segment from a field measurement and one poor quality

[7] The quantity and type of practice segments presented to the listener panels for the IS-95-based CDMA (TAG 2) testing (as well as the other subsequent JTC PCS technologies) was different from that presented to the listener panels for the PCS 1900 (TAG 5) testing. More practice segments were used and the use of more 64-kbps wireline voice segments was initiated.

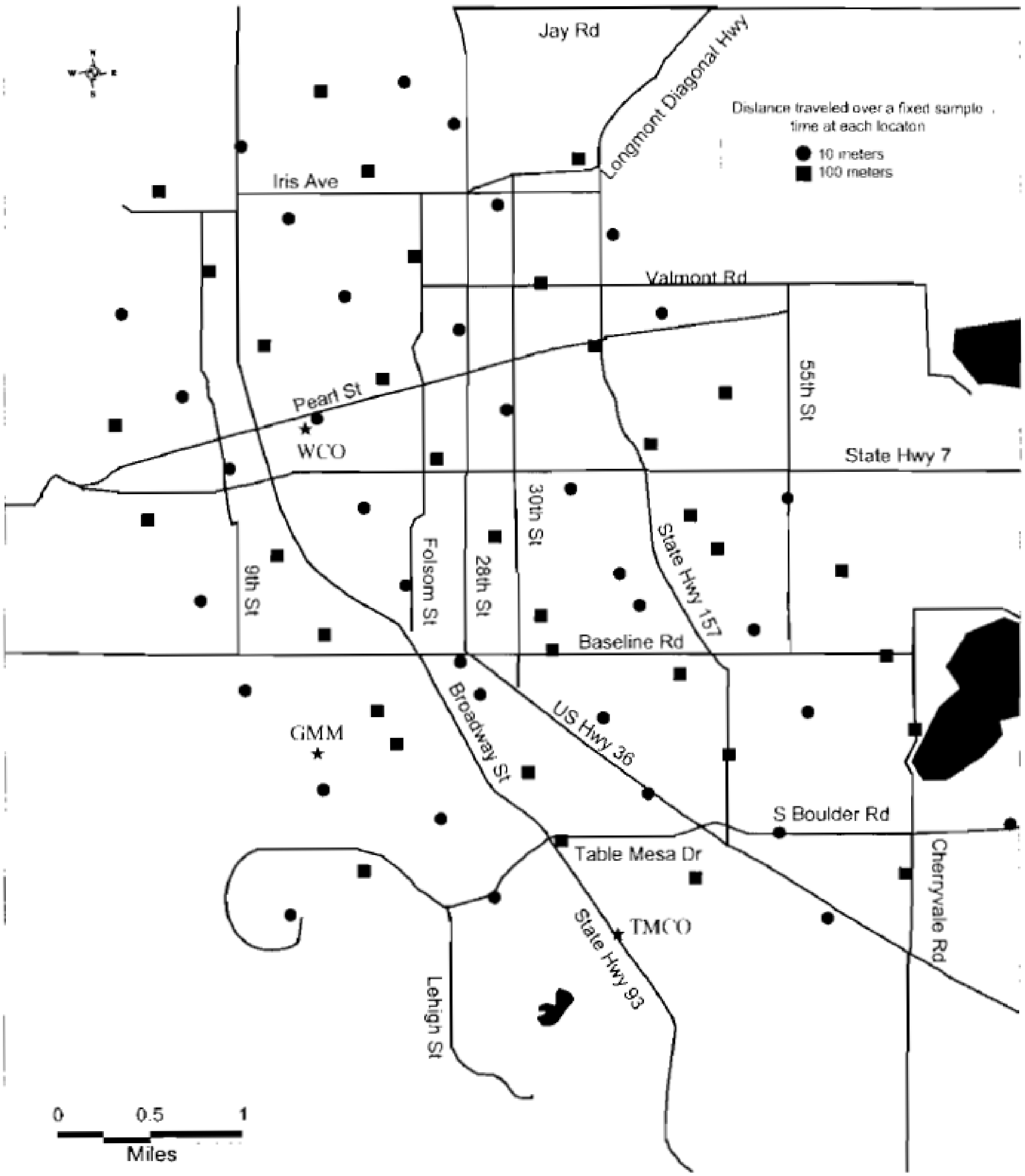

Figure 4.14. Quasi-stationary measurement locations and vehicular speed used at each location for TAG 2.

voice segment from a field measurement. Each of these wireline segments was presented twice. The codec segments used in training were either full-rate or half-rate segments, with an average FER of either 0%, 1%, 3%, or 5%. The goal was to demonstrate a wide range of quality in the practice segments presented to listeners. It was felt that the voice segments collected during the quasi-stationary measurements did not represent the full range of quality needed, therefore the codec segments were recorded in the lab. The practice segments allowed inexperienced listeners to gain exposure to the full range of possible voice quality, from excellent to poor, and ensured proper scaling of the MOS's.

After the practice segments were presented, the subjects in each of the four groups were asked to rate 33 voice segments: 3 reference voice segments and 30 voice segments from the field trial measurements. Every listener within a group listened to the same voice segments. The reference voice segments were either 64-kbps wireline segments or poor quality voice segments collected from the field measurements. The 30 voice segments from the field trial measurements came from the uplink or downlink measurements at 60 out of the 69 measurement locations. Only 60 of these locations were used in the MOS testing to limit the number of voice segments each listener had to rate. As stated earlier, each of the voice segments from the field trial measurements consisted of 10 male and 10 female sentences. For the MOS testing, the first sentence and the last sentence of each recorded voice segment were not used. This was done to avoid the possible inclusion of sentences that had been cut short due to starting a recording too late or stopping a recording too soon. Therefore, during MOS testing, a total of 18 sentences (out of the original 20 sentences) were presented to listeners for each segment. For each segment, nine male sentences and nine female sentences were presented. The nine male sentences were presented before the nine female sentences half of the time. In addition, the order of presentation of the voice segments was randomized. Subjects were given a 15- to 20-min break half way through the session. After all segments were presented, subjects filled out a post-trial questionnaire.

For each voice segment, voice quality ratings (answers to the question "How would you rate the overall quality of the sound?") from each subject within a group were averaged to obtain an MOS. The results from all four of the groups (a total of 120 voice segments) are shown in the histogram in Figure 4.15. Overall, the voice segments were rated favorably, with 88% of the segments rated between fair and excellent. The average MOS was 3.56 and the standard deviation was 0.50.

Figures 4.16 and 4.17, show histograms of MOS's for the uplink and downlink, respectively. The average MOS for the uplink was 3.65 and for the downlink was 3.47. A t-test revealed that there was a statistically significant difference in the average MOS's between the uplink and downlink.

The relationship between the MOS's and some of the objective measures was initially investigated by generating some scatter plots. Figure 4.18 shows the relationship between the MOS's and the average downlink RSS. For each MOS, the average RSS is the average of the measured RSS values computed over the duration of the voice segment. In Figure 4.18, a large variation in MOS's is seen for all values of average RSS. Figures 4.19 and 4.20 show the relationship between the MOS's and the average FER for both the uplink and downlink,

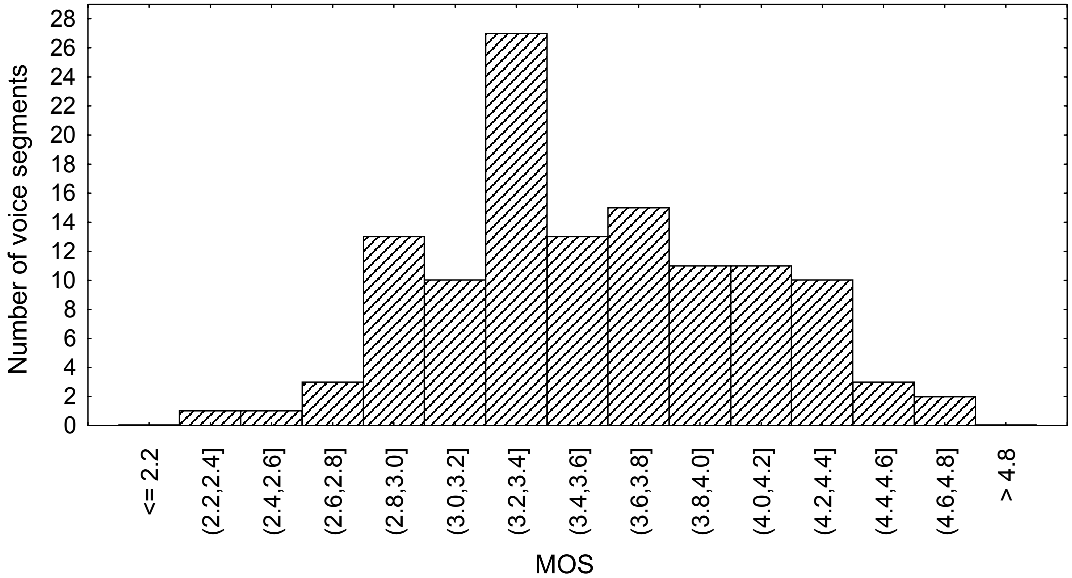

Figure 4.15. Histogram of mean opinion scores (MOS's) for all voice segments (TAG 2).

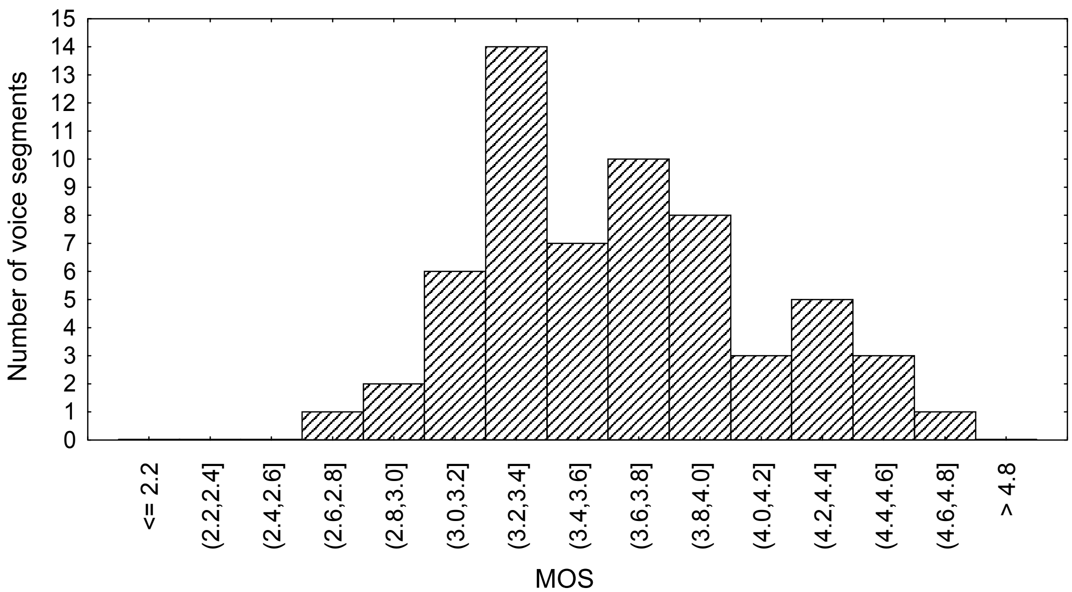

Figure 4.16. Histogram of mean opinion scores (MOS's) for the uplink (TAG 2).

respectively. As with the average RSS, for each MOS, the average FER is the average of the measured FER values computed over the duration of the voice segment. A large variation in MOS's for all average FER's on both the uplink and downlink is seen in Figures 4.19 and 4.20. The RSS and the FER, when averaged over the entire voice segment, do not appear to be accurate predictors of listener satisfaction.

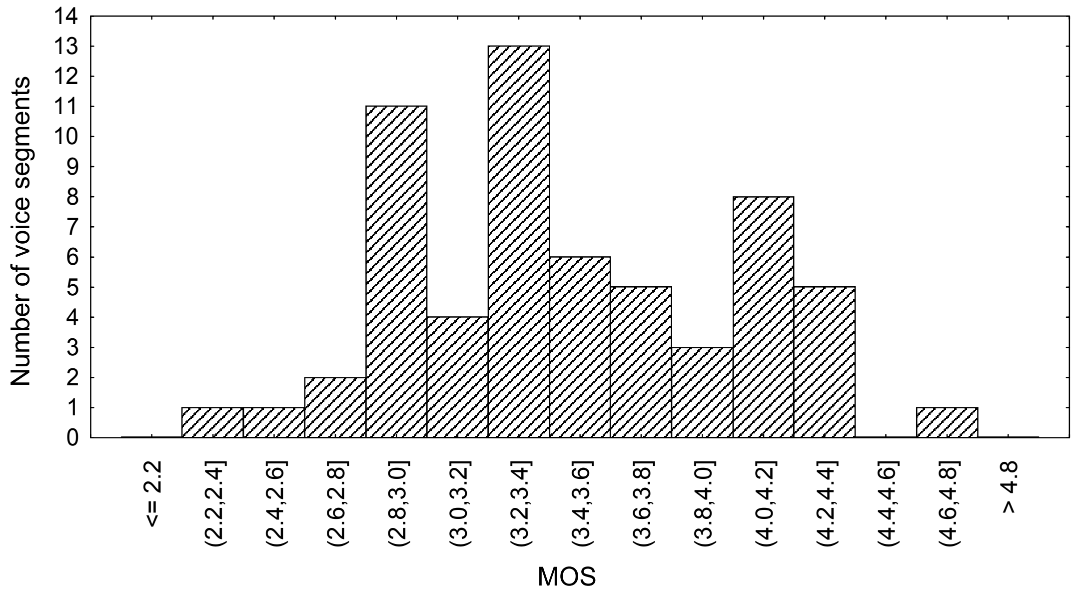

Figure 4.17. Histogram of mean opinion scores (MOS's) for the downlink (TAG 2).

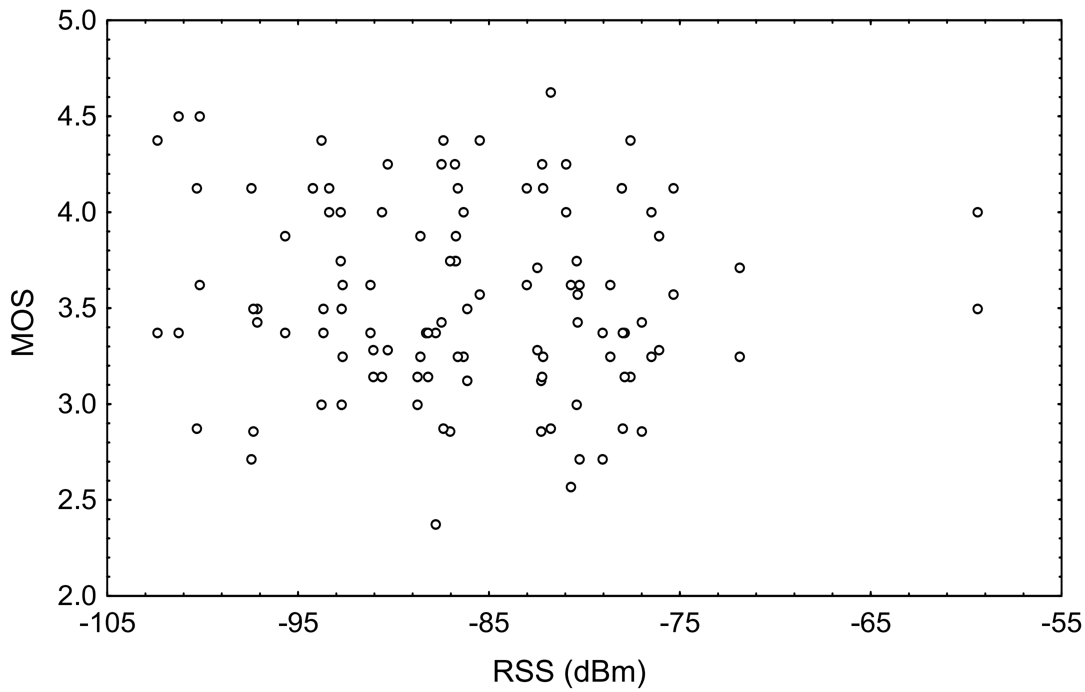

Figure 4.18. Mean opinion score (MOS) vs. average downlink received signal strength (RSS; TAG 2).

Pearson product-moment correlations were performed to determine the correlation between MOS's and average RSS, MOS's and average FER, and MOS's and other objective measures. The correlation coefficient between MOS and average RSS was -0.08 and that between MOS and average FER was 0.01. The highest correlation coefficient between MOS and any objective measure, that between MOS and average mobile transmit gain adjust, was only 0.15. These very low correlations between MOS and averaged objective measures suggest that a linear relationship between MOS and the objective measures does not exist.

While there does not appear to be linear relationship between MOS and the objective measures, there still may be a consistently increasing or decreasing relationship between them. The Spearman rank correlation can be used to determine if a consistently increasing or decreasing trend may exist between MOS and the objective measures. Spearman rank correlations were performed to determine the correlation between the ranks of MOS and the ranks of average downlink RSS, between the ranks of MOS and the ranks of average FER, and between the ranks of MOS and the ranks of other objective measures. The Spearman rank correlation coefficient between MOS and average downlink RSS was -0.08 and that between MOS and average FER was 0.05. The highest Spearman rank correlation coefficient between MOS and any objective measure, that between MOS and average QTSO target E_b/N_o, was only 0.16. These very low rank correlations between MOS and averaged objective measures suggest that a consistently increasing or decreasing relationship between MOS and the objective measures does not exist.

Note that the objective measures were averaged over the entire length of the voice segment. By analyzing the instantaneous variation or possibly minimum and maximum values of the objective measures within the voice segment, further insight may be gained on the behavior of MOS's.

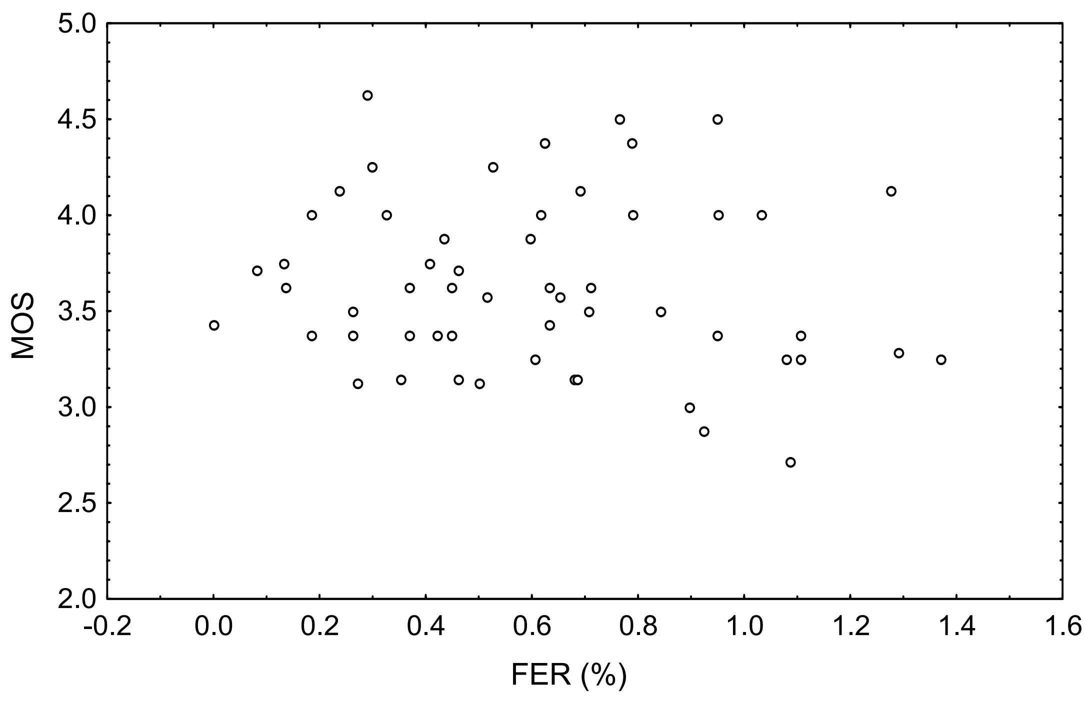

Figure 4.19. Mean opinion score (MOS) vs. uplink frame error rate (FER; TAG 2).

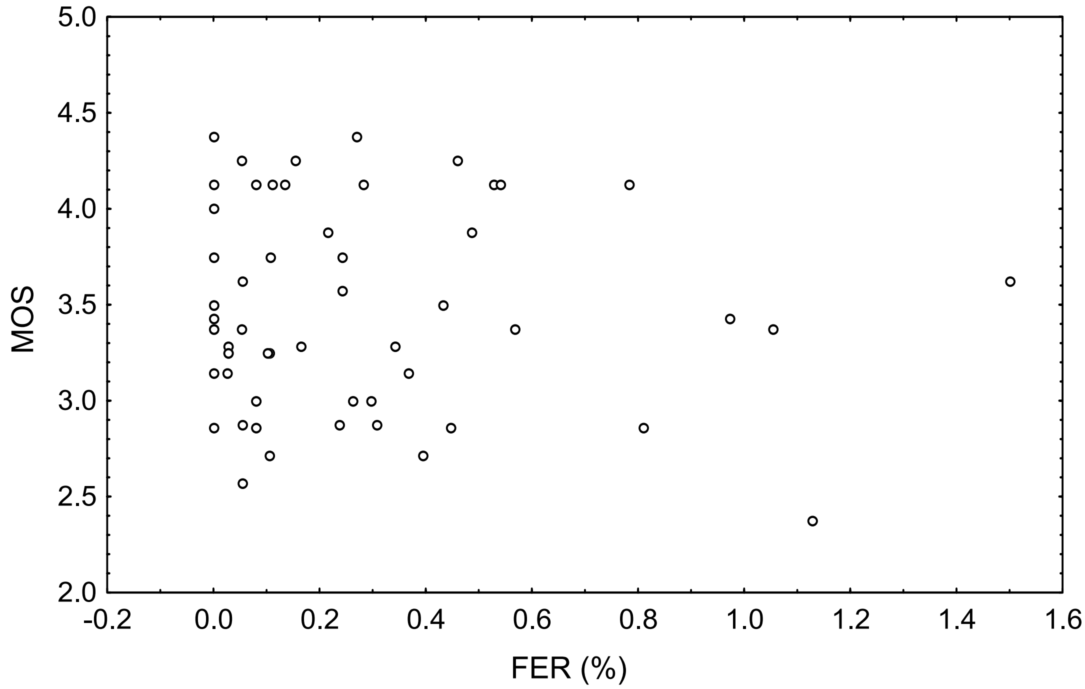

Figure 4.20. Mean opinion score (MOS) vs. downlink frame error rate (FER; TAG 2).

By gathering listeners' comments from post-test questionnaires, more information about the nature of MOS's was obtained. Namely, it is evident from questionnaires that there are several types of distortions in quality possible in the voice recordings of the IS-95-based CDMA system according to listeners:

1) echo;
2) fading/muting;
3) synthesized voice;
4) "chirping," or "squeaking" background noise; and
5) a clearer male (vs. female) voice.

The nature of these distortions are likely judged differently by different listeners. Intelligibility and speaker recognition are two main aspects of perceived quality. In the case of most of the voice segments for the IS-95-based system, intelligibility remained high. The above-mentioned types of distortions seem to affect a listener's ability to recognize the speaker more than their ability to understand the speech. As a result, overall MOS's were high. However, it is of interest to note that the MOS's were significantly lower in one particular group of listeners with average MOS's of 2.95 vs. the average of 3.56 over all MOS's.

There were more women aged 35 and above in this group than any other. Compared to other age/gender groups, marked differences in MOS's for this demographic group were seen. Two of the women in this group stated in their post-test questionnaires that the female voice was more degraded than the male voice in many of the voice segments. When asked whether their ratings of acceptability would have changed if told to consider them as wireline samples vs. portable, the answer was no for all four women over 35 in this group. One hypothesis for explaining the differences in MOS's for this demographic group is that female listeners are harder "graders" of quality than others. Another hypothesis is that female listeners are more sensitive to distortions in higher frequencies, and the voice segments reflected such a distortion. Further research is necessary to pinpoint whether there is a significant difference between ratings given by females over the age of 35 and those given by others, and to pinpoint potential reasons for this. It is still unclear at this stage why some listeners were harder "graders" than others.

The variability in scoring may indicate subjects graded voice segments with unrealistic criteria. Many of the voice segments were nearly identical in quality, and yet judged quite differently by the same subject. This might indicate randomness in scoring among listeners. Further analysis may pinpoint the likelihood of random or unrealistic scoring.

4.6.3 Expert Listener Assessment

In addition to being rated by listener panels in MOS testing, the voice segments were rated by an expert listener. The expert listener ratings followed the identical procedure as in the PCS 1900 (TAG 5) testing. This procedure is described in Section 3.6.3.

Figure 4.21 shows the relationship between expert listener ratings and percent acceptability (the percentage of listeners rating a given voice segment as acceptable). The boxes represent the middle half of the data (from the 25th percentile to the 75th percentile). The solid circles represent the median percent acceptabilities for each of the expert listener ratings. The lines extending out of the boxes depict the spread of the data.

For the most part, the definitely acceptable ratings of the expert listener were relatively good indicators of percent acceptability. In the definitely acceptable category, most of the voice segments were rated as acceptable by 70-100% of the subjects; a few voice segments were rated as acceptable by less than 70% of the subjects. The mean percent acceptability in the definitely acceptable category was 90.2%, well within the target range of 70-100%. The marginally acceptable and unacceptable expert listener categories show more variance. For the voice segments rated as marginally acceptable by the expert listener, the mean percent acceptability was 71.9%, slightly above the target range of 30-70%. The mean percent acceptability associated with voice segments rated as unacceptable by the expert listener was 59.5%, well above the target 30% and below.

The Pearson product-moment correlation coefficient between MOS and percent acceptability was 0.67. This indicates some correlation between these quantities as would be expected. The Pearson product-moment correlation coefficient between MOS and expert listener rating was 0.40, indicating that a strong linear relationship between these measures does not exist.

It is possible that expert listener rating, downlink RSS, FER, and the other objective measures including transmit gain adjust, mobile transmit power, mobile received pilot E_c/I_o, QTSO target E_b/N_o, and QTSO forward power control gain can be predictors of MOS when all are taken together. A multiple regression analysis was completed in order to determine if MOS is

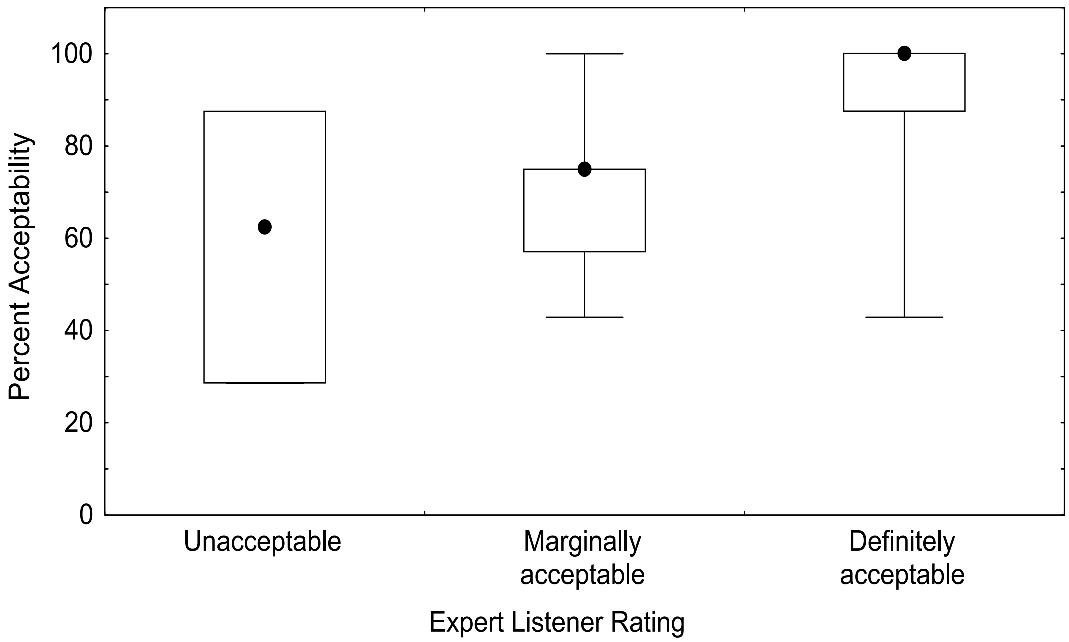

Figure 4.21. Percent acceptability vs. expert listener rating (TAG 2).

related to a combination of all the above factors. Interactions among the various factors were not considered in the analysis.

The result showed that only 21% of the variance can be explained by expert listener rating, downlink RSS, FER, and the other objective measures listed above. The result of the multiple regression analysis also showed that all these factors reliably account for the 21% of variance. From this analysis, it appears that at least at first glance, the expert listener rating, downlink RSS, FER, and the other objective measures listed above taken together do not offer any substantial predictive power over MOS's.

4.6.4 Voice Quality Handoff Measurements

Continuous voice recordings were made as the mobile unit traveled along routes through handoff areas. While the measurement van was in motion, an audio source tape was transmitted over the uplink and downlink simultaneously. The source tape consisted of Harvard sentences and was played continuously as the measurement van traveled along each route. The received voice transmissions were recorded on digital audio tape at the receiver for the uplink and at the receiver for the downlink. Continuous voice recordings were made along the route until the call was finally dropped. The routes were selected according to previous tests at the BITB. One route was driven along Broadway, at a vehicular speed of 25-30 mph. The second route was driven along Foothills Parkway, at a vehicular speed of 55-60 mph. The beginning of a route was selected within the coverage area. Two runs were conducted along each route, one traveling from the north to the south, and the other in the opposite direction. Voice quality was assessed using the expert listener methodology described in Section 3.6.3.

For the voice quality handoff testing, an expert listener rating was made for each 4-s period of the continuous voice recordings taken along a measurement route. This equates to roughly one rating per sentence.

Handoffs were indistinguishable to the expert listener, including soft, softer, and soft-softer handoffs. In general, voice quality remained good until the coverage boundaries were approached. Then, voice quality dropped quickly, most likely due to dramatic changes in terrain.

4.7 Manufacturer's Statement

The statement provided by Qualcomm Incorporated is included in this section. This statement is identical to that given in [2], except for some minor editorial changes.

Qualcomm Incorporated would like to thank U S West and ITS for their support during the demonstration of the performance of the U S CDMA PCS radio air interface at 1.9 GHz at the BITB in Colorado. The test bed is operated by U S West. In particular, the support provided by M. Laflin of ITS and the technical staff of U S West is acknowledged and very much appreciated. Qualcomm would further like to thank all TAG 2 participants for their hard work in completing the ANSI J-STD-008, personal station-base station compatibility requirements for 1.8- to 2.0-GHz CDMA PCS, and for their support of this test. Funding for the test was provided by Qualcomm, AT&T, Motorola, Nokia, and Nortel.

Qualcomm was pleased to supply equipment to demonstrate the CDMA PCS air interface in the BITB as part of the standardization requirements of the JTC. TAG 2 of the JTC developed the PCS standard which is based on TIA/EIA dual mode cellular CDMA standard IS-95-A. The resulting ANSI PCS standard has been approved for publication by both T1P1 and TR46.

The CDMA system performed quite well as the results of the testing show. The results are consistent with trials of the CDMA PCS technology in other locations. CDMA provided consistently high voice quality and excellent coverage. The test was conducted with a loading equivalent to 10 calls in every sector of each cell using a single CDMA radio channel.

Because of time constraints, there was no attempt to tune or optimize the system used in Boulder, Colorado. As a result, the percentage of soft handoff is higher than would be targeted in an operational deployment. An interesting test would have been to investigate the influence of the handoff parameters and higher load on the percentage of soft handoffs and overall system performance.

The roving test system (RTS) used for the JTC tests is a noncommercial test system that employs older 3-chip CDMA technology in the base station. The more recently available single chip cell-site modem (CSM) that is being employed in operational CDMA systems requires about 2 dB less E_b/N_o on the reverse traffic channel (uplink) — the exact amount depends upon the multipath environment. In addition, in the BITB tests, the 8-kbps IS-96-A speech codec was used. Higher voice quality is provided in deployed CDMA systems tests that use the more recently available 13-kbps speech codec.

5. TAG 4 (IS-136-BASED SYSTEM) TESTING

This section describes the test plan, methodology, and results for the technology field trial conducted by TAG 4. The technology tested was the PCS variant of the IS-136 cellular TDMA standard. The base station equipment was provided by AT&T Network Wireless Systems and the mobile units were provided by Ericsson and Nokia. The TAG 4 field tests examined area coverage, handoff, and voice quality; and the effects of co-channel and adjacent channel interference on system performance.

The information presented in this section is taken from [4]. The reader is referred to [4] for a more complete and detailed presentation of the TAG 4 technology field testing at the BITB.

5.1 TAG 4 Test System Configuration

The block diagram of the test system configuration is shown in Figure 5.1. The test system consisted of the three cell sites in Boulder, Colorado (WCO, TMCO, and GMM) and one mobile switching center (MSC) supplied by AT&T. Plain old telephone service (POTS) was provided remotely by an AT&T switch located in Whippany, New Jersey. The MSC was also located in Whippany, New Jersey. The MSC was connected to the WCO cell site via a T1 circuit. The T1 circuit connects to the MSC by using a digital automatic cross connect system (DACS) and two T1 circuits. Connection was made to the WCO base station by using a micro DACS (μDACS). Another T1 circuit connects the MSC to the TMCO base station.

The TAG 4 system that was tested during the field trials was based on an IS-136 cellular A-band system upbanded to PCS D-band frequencies. An analog control channel was used in accordance with the IS-136 specification. The base station and mobile unit transmit frequencies in the PCS-D band are shown in Table 5.1.

5.2 Calibration

A calibration of the base station RSS was performed by injecting a digitally modulated signal of known level into the receiver and then generating a table of scale factors used by the logging software to provide accurate RSS reports during the field trial. This was done for all active sectors at each cell site.

The input signal into the base station receiver was provided by a digital modulation signal generator. Losses in the cabling between the signal generator and the input of the base station receiver were measured and a correction, or offset, was added to the signal generator output-level reading. The level of the input signal was varied over the dynamic range of the receiver to generate the table of scale factors. Note that for this calibration procedure, the diversity branch of the receiver was terminated with a 50 Ω load.

The mobile units (provided by Ericsson and Nokia) were checked for accuracy of the reported RSS values. This was done by injecting a digitally modulated signal of known strength directly

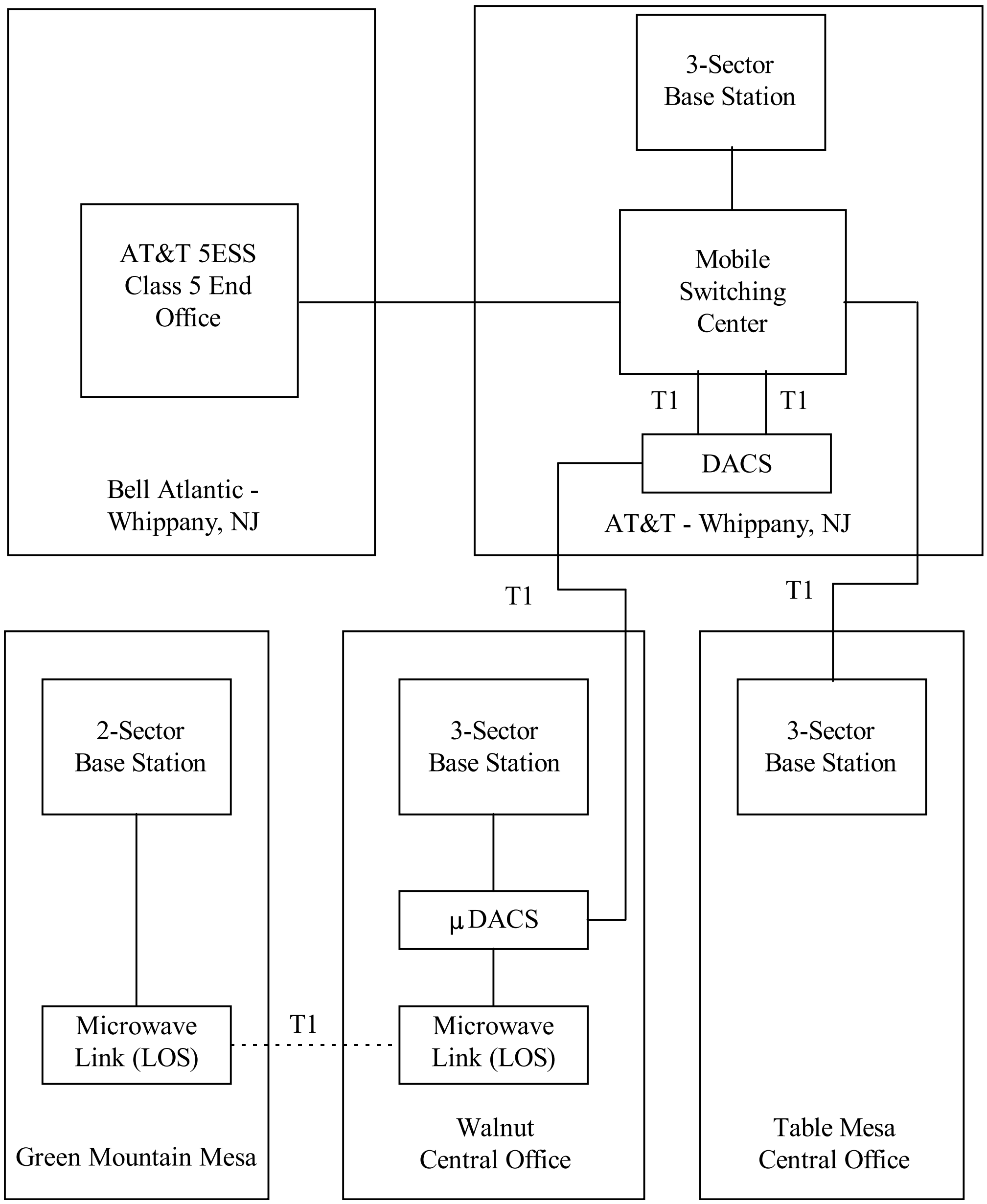

Figure 5.1. Block diagram of the TAG 4 test system configuration.

into the antenna port of the mobile unit. The level of the input signal was then compared to the RSS value reported by the mobile unit. All of the mobile units had reported values of RSS within 4 dB of the actual input signal level. Note that both the Ericsson and Nokia mobile units report RSS in 2-dB steps.

5.3 Area Coverage Testing

To provide maximum coverage area, both the base station and the mobile unit used full power outputs. The base station power for each cell site was set so that the output power was approximately 50 W ERP. Measurements to show area coverage were taken with the mobile

Table 5.1. TAG 4 Base Station And Mobile Unit Transmit Frequencies

Cell Site	Sector	Type of Channel	PCS Channel #	Mobile Unit Transmit Frequency (MHz)	Base Station Transmit Frequency (MHz)
WCO	North	Control	632	1868.94	1948.98
WCO	North	Traffic	532	1865.94	1945.98
WCO	Southeast	Control	635	1869.03	1949.07
WCO	Southeast	Traffic	535	1866.03	1946.07
WCO	Southwest	Control	628	1868.82	1948.86
WCO	Southwest	Traffic	528	1865.82	1945.86
GMM	North	Control	644	1869.30	1949.34
GMM	North	Traffic	544	1866.30	1946.34
GMM	Southeast	Control	637	1869.09	1949.13
GMM	Southeast	Traffic	537	1866.09	1946.13
TMCO	North	Control	625	1868.73	1948.77
TMCO	North	Traffic	525	1865.73	1945.77
TMCO	Southeast	Control	639	1869.15	1949.19
TMCO	Southeast	Traffic	539	1866.15	1946.19
TMCO	Southwest	Control	642	1869.24	1949.28
TMCO	Southwest	Traffic	542	1866.24	1946.28

unit[8] located in a mini-van as in both the TAG 5 and TAG 2 testing. The mobile unit was mounted inside the van on the same wooden structure used in all of the JTC PCS technology field trials. This structure is described in Section 3.3. The measurements were taken by driving along routes (radials) away from the cell site. When time permitted, measurements along additional routes in between the radials were taken.

Only one cell site was activated at a time during area coverage testing; all other cell sites were powered down. Handoff was allowed between the sectors of the active cell. The data were collected both at the mobile unit and at the base station. The data collected at the mobile unit included GPS location, velocity, and time; downlink RSS; downlink BER class; and downlink FER (or number of frame errors); in addition to other system parameters. BER class is a measure of the average BER and is related to the average BER as shown in Table 5.2. The data collected at the base station included GPS time, uplink RSS, uplink FER (actually number of frame errors), uplink number of bits in error per second, and downlink RSS and BER class (as reported by the mobile unit), in addition to other system parameters.

Calls were originated from the mobile unit prior to the start of data collection. Collection of mobile data and base station data was initiated as the measurement van began traveling away from the cell site along the drive route. In some cases, shadowing or poor coverage would cause the call to be dropped along the route; in those cases, the call was reinstated if the RSS improved in a short distance. At the end of the route, the data collection was stopped and the data were saved to disk. All data collected were averaged over 1 s. Because of mobile unit sampling rate problems, the data reported by the mobile unit to the base station (and recorded at the base station) were used in the data analysis instead of the data recorded at the mobile unit. Therefore, the downlink FER was not available for the data analysis; downlink BER class was used instead.

[8] Measurements were also taken with an antenna mounted on the roof outside of the mini-van. For purposes of brevity and because it was not a formal part of the JTC PCS technology field trials, TAG 4 measurements performed with the antenna located out of the measurement van are not discussed in this NTIA report.

Table 5.2. Relationship Between Bit Error Rate (BER) Class and Average BER

BER Class	Average BER
0	less than 0.01%
1	0.01% to less than 0.1%
2	0.1% to less than 0.5%
3	0.5% to less than 1.0%
4	1.0% to less than 2.0%
5	2.0% to less than 4.0%
6	4.0% to less than 8.0%
7	greater than 8.0%

5.3.1 TMCO Area Coverage Data

The test procedure followed and data collection methodology used are explained in Section 5.3. All three sectors were active for this cell site. The mobile units used for the TMCO cell were manufactured by Ericsson.

Downlink RSS as a function of distance is shown in Figure 5.2. The data included in this figure are the overall data for this cell, excluding RSS values less than -109 dBm. RSS values less than -109 dBm were excluded because -109 dBm is the threshold level for normal system operation of the mobile unit receiver. The large variation in RSS seen in Figure 5.2 is due to the irregularity of the terrain. Note that relatively strong signals (approximately -72 to -80 dBm) exist far away from the cell (approximately 12.5 mi). Those signals were recorded in areas having LOS propagation between the mobile unit and the base station.

A rough estimate of the coverage area was determined by assuming that an RSS of -100 dBm or greater is desired. The measured RSS data along all of the routes driven within the cell were used to determine the coverage area. Due to the irregularity of the terrain, the RSS varied significantly along the TMCO routes, crossing the -100 dBm level several times before finally staying below -100 dBm. The point along each route where the RSS first dropped below -100 dBm was used to define the coverage boundaries. For this case, the coverage boundaries were approximately 8.6 mi north-northeast, 3.2 mi to the northeast, 4.8 mi due east, and 2.71 mi due south. Due to the higher elevation of this site (relative to the WCO cell), this cell had greater coverage in the northern-northeastern direction than did the WCO cell.

Figure 5.3 shows the histogram of downlink RSS values. Again, RSS values less than -109 dBm were excluded. The mean RSS is -88.2 dBm with a standard deviation of 10.0 dB. Figure 5.4 shows a histogram of the uplink RSS values for the TMCO site.

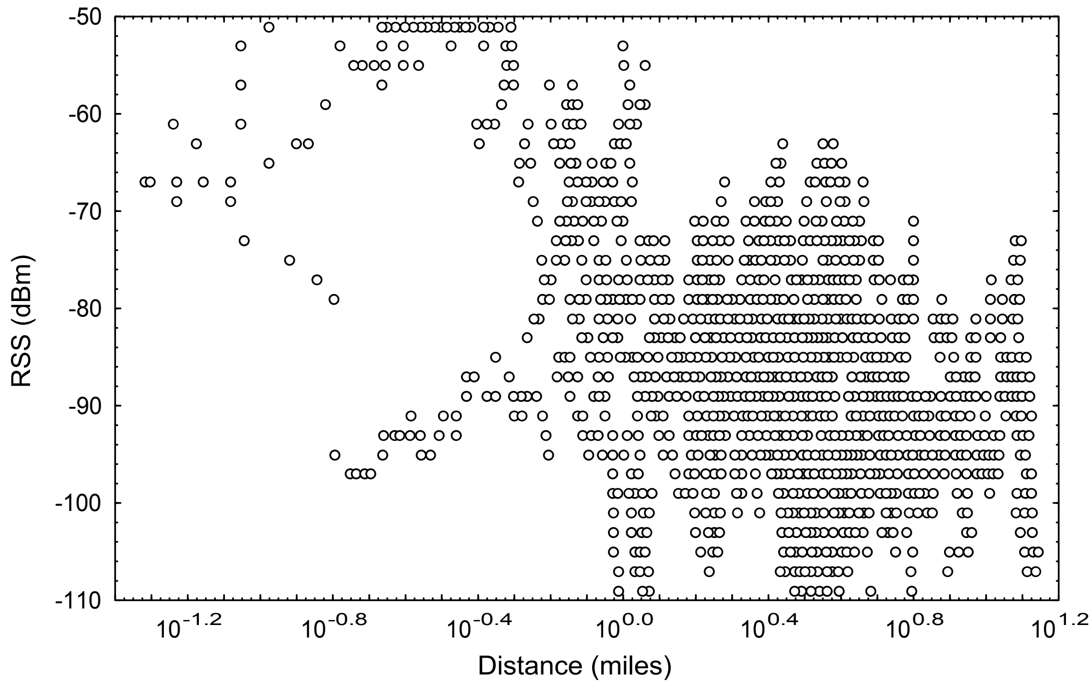

Figure 5.2. Downlink received signal strength (RSS) vs. distance (TAG 4, TMCO cell).

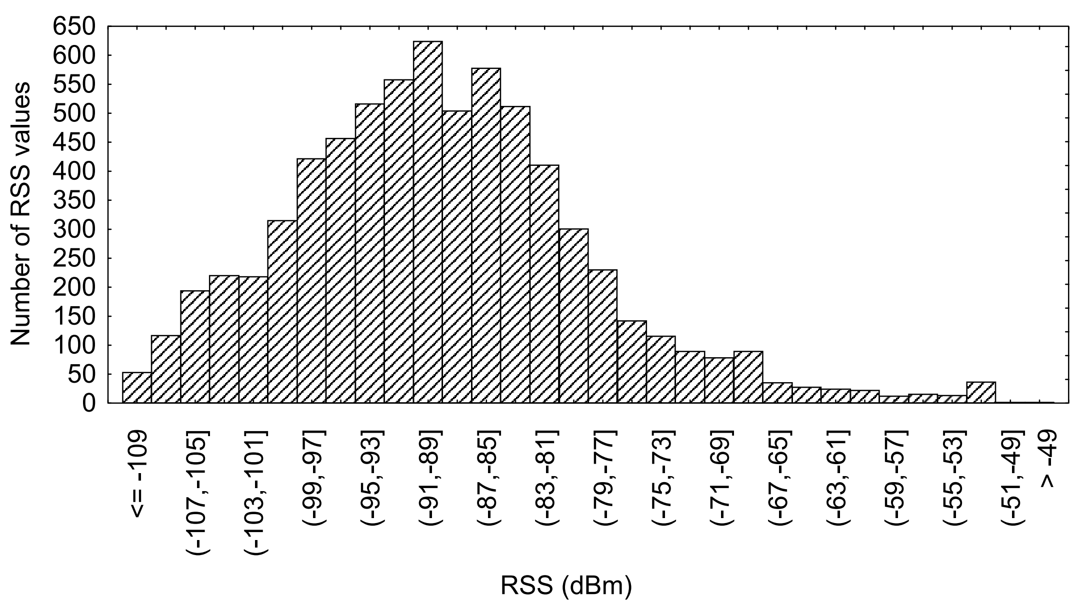

Figure 5.3. Histogram of downlink received signal strength (RSS) values (TAG 4, TMCO cell).

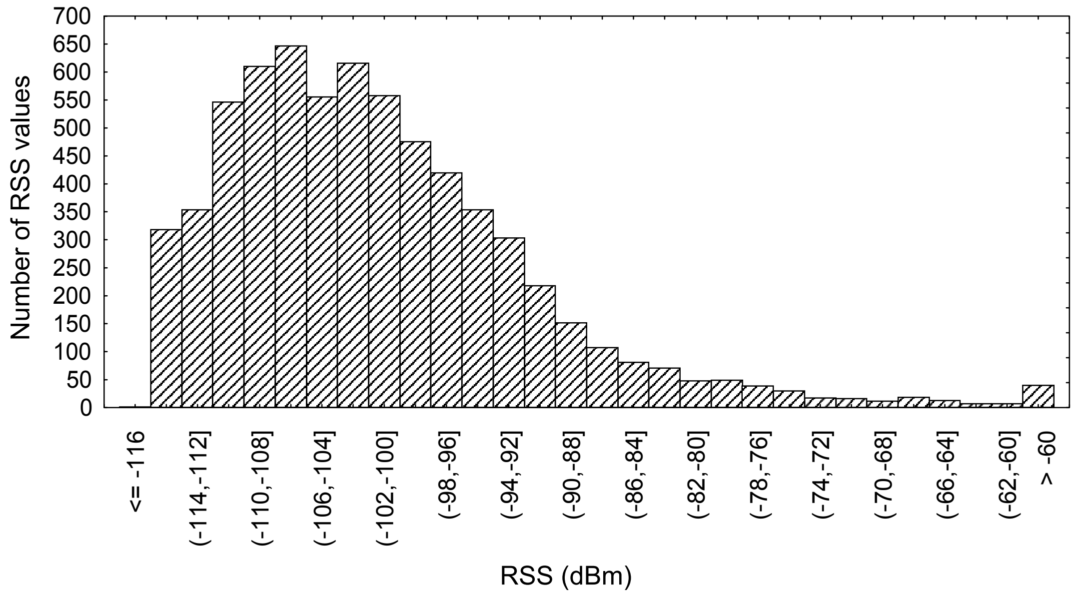

Figure 5.4. Histogram of uplink received signal strength (RSS) values (TAG 4, TMCO cell).

Link balance between the uplink and the downlink was analyzed by generating a scatter plot of downlink RSS vs. uplink RSS. Linear regression was then performed to determine the best fit for the data. Figure 5.5 shows the scatter plot and the results of the linear regression. Both the scatter plot and the linear regression indicate that the link was unbalanced.

To estimate the amount of link imbalance, a histogram of the difference between the downlink and uplink RSS was generated. This histogram is shown in Figure 5.6. The mean difference between downlink and uplink RSS was 13.3 dB with a standard deviation of 4.36 dB. The histogram indicates that the system was uplink-limited (i.e., RSS at the base station was typically weaker than RSS at the mobile unit). This is probably the result of operating at full power to provide wide area coverage.

The uplink FER as a function of uplink RSS is shown in Figure 5.7. As expected, the FER decreased as the RSS increased. The FER was below 1% for RSS values of -106 dBm and higher.

5.3.2 WCO Area Coverage Data

As for the TMCO cell, the test procedure followed and data collection methodology used are explained in Section 5.3. All three sectors were active for this cell site with handoff allowed between the sectors. About half of the measurement routes in the WCO cell used mobile units manufactured by Ericsson; the other half used mobile units manufactured by Nokia. Since there were some regions of overlapping of the different measurement routes, data were collected in some areas by both the Nokia and Ericsson mobile units.

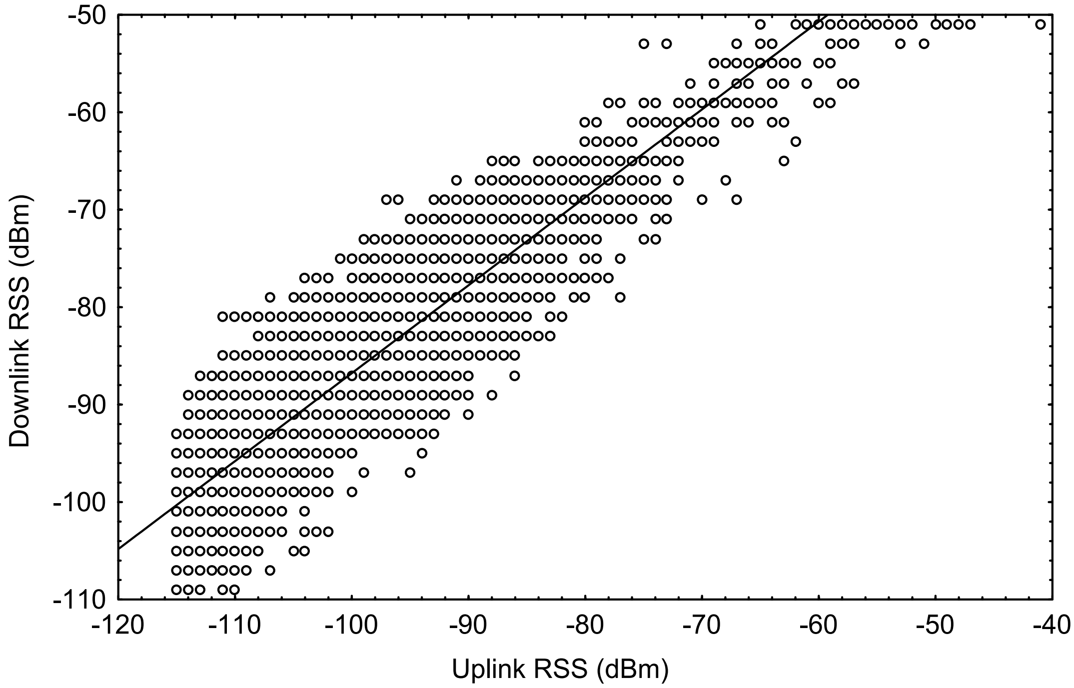

Figure 5.5. Downlink received signal strength (RSS) vs. uplink RSS (TAG 4, TMCO cell).

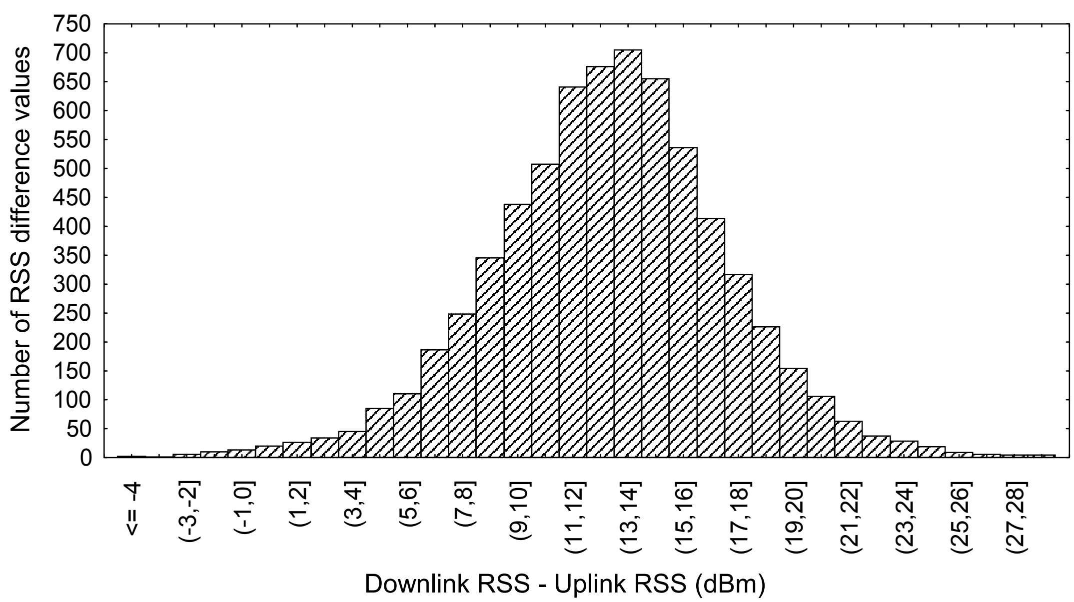

Figure 5.6. Histogram of difference between downlink and uplink received signal strength (RSS; TAG 4, TMCO cell).

Plots of downlink RSS as a function of distance for the Ericsson and Nokia mobile units are shown in Figures 5.8 and 5.9, respectively. The data included in these figures are the overall

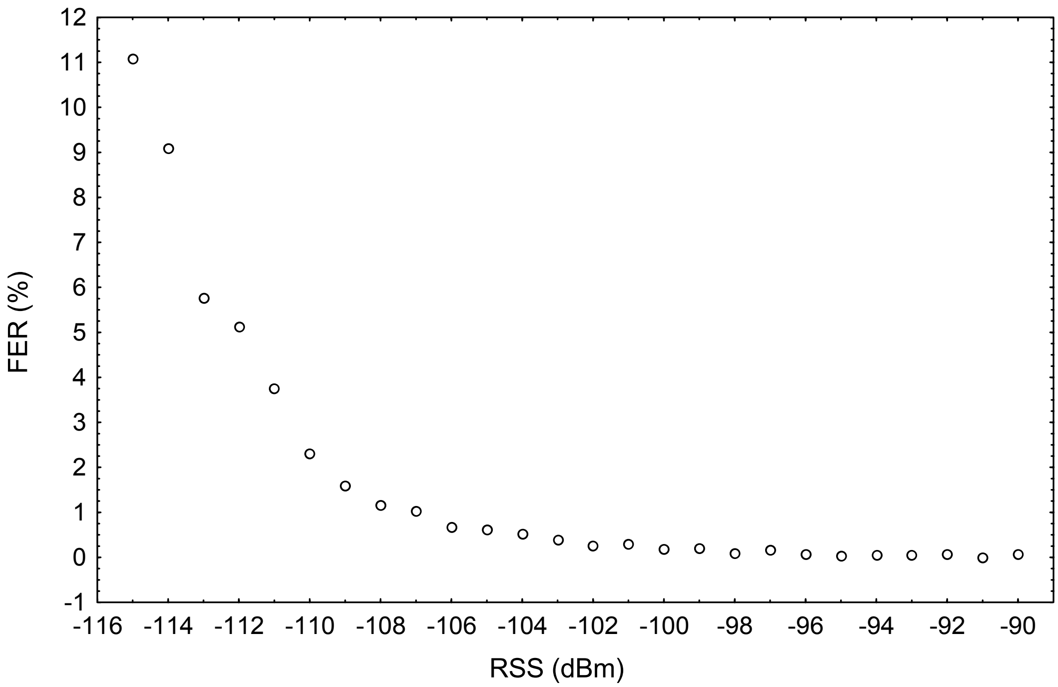

Figure 5.7. Uplink frame error rate (FER) vs. received signal strength (RSS; TAG 4, TMCO cell).

data for this cell, excluding RSS values less than -109 dBm. RSS values less than -109 dBm were excluded because -109 dBm is the threshold level for normal system operation of the mobile unit receiver. As in the TMCO cell, there is a large variation in the RSS for a given distance. The large variation in RSS seen in Figures 5.8 and 5.9 is due to shadowing and the particular choice of routes. Note that some signals with an RSS greater than -100 dBm exist 9 mi away from the center of the cell. Comparing this maximum distance with the maximum distance obtained for the TMCO cell (approximately 12.5 mi) suggests that the WCO cell has a smaller coverage area than the TMCO cell.

A rough estimate of the coverage area was determined by assuming that an RSS of -100 dBm or greater is desired. The measured RSS data using both the Ericsson and Nokia mobile units along all of the routes driven within the cell were used to determine the coverage area. The coverage boundaries were approximately 1.26 mi due north, 3.2 mi due east, 2.0 mi to the southeast, and 2.0 mi to the west. Once the RSS dropped below -100 dBm in these directions, it generally stayed below that value. Due to the lower elevation of this site (relative to the TMCO cell) and the terrain profile, this cell had a smaller coverage area than did the TMCO cell.

Figure 5.10 shows a histogram of the uplink RSS values for the WCO site. The mean uplink RSS was -99.49 dBm with a standard deviation of 9.94 dB.

Link balance between the uplink and the downlink was analyzed by generating histograms of the difference between the downlink and uplink RSS for both the Ericsson and Nokia mobile units. These histograms are shown in Figures 5.11 and 5.12. The mean difference between

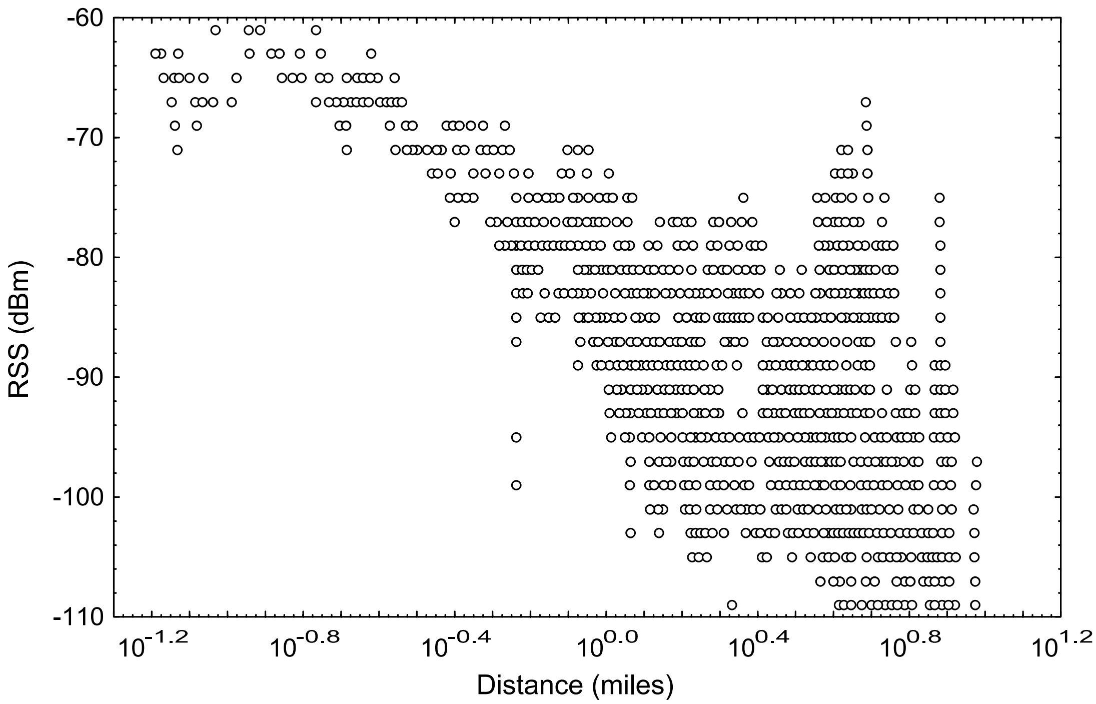

Figure 5.8. Downlink received signal strength (RSS) vs. distance (TAG 4, WCO cell, Ericsson mobile unit).

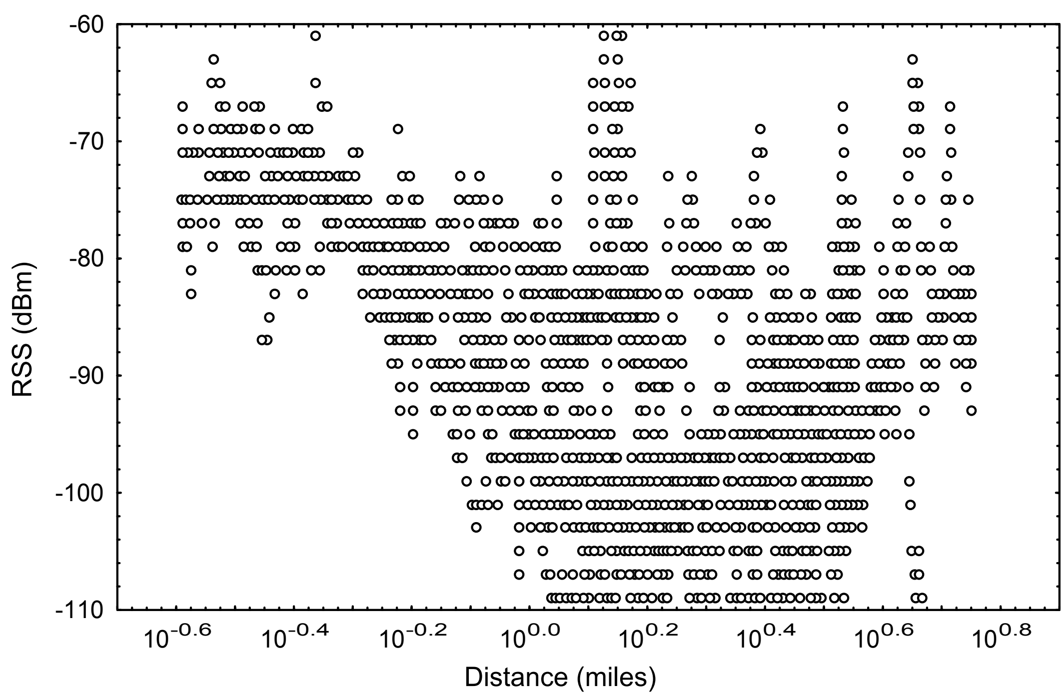

Figure 5.9. Downlink received signal strength (RSS) vs. distance (TAG 4, WCO cell, Nokia mobile unit).

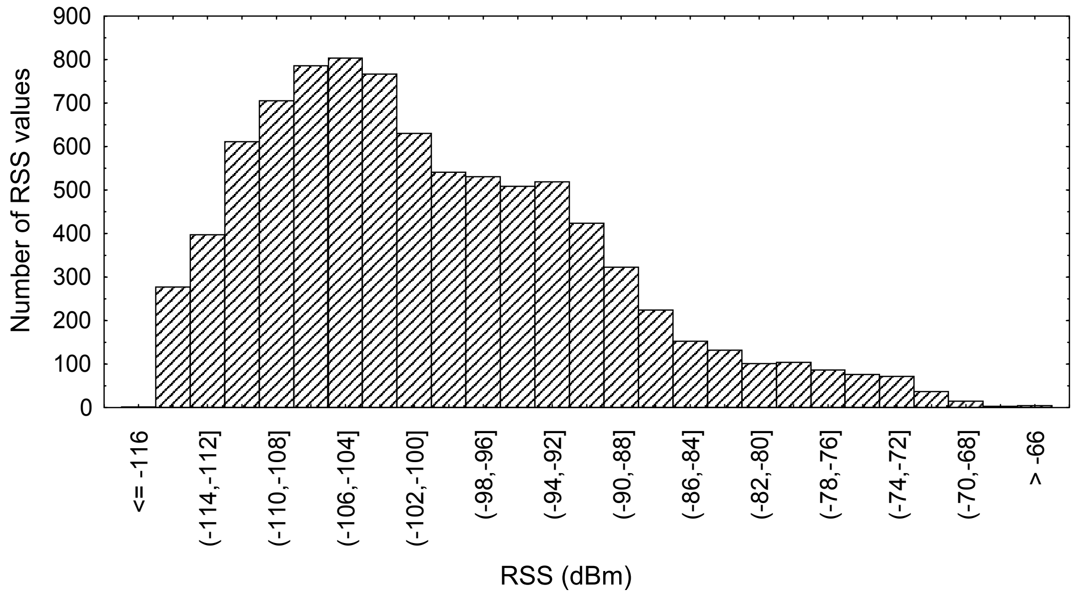

Figure 5.10. Histogram of uplink received signal strength (RSS) values (TAG 4, WCO cell).

downlink and uplink RSS for the Ericsson mobile unit was 13.09 dB with a standard deviation of 4.38 dB. The mean difference between downlink and uplink RSS for the Nokia mobile unit was 6.77 dB with a standard deviation of 3.98 dB. These histograms indicate that the link was unbalanced and that the system was uplink limited (i.e., RSS at the base station was typically weaker than RSS at the mobile unit). This is probably the result of operating at full power to provide wide area coverage.

The difference in link imbalance between the Ericsson and Nokia mobile units may have been caused by differences in the RSS reporting accuracy (discussed in Section 5.2), the antenna frequency response of the Nokia handset antenna (higher gain at the mobile transmit frequency than at the mobile receive frequency), and the higher transmit power of the Nokia mobile unit (approximately 0.5 dB).

The uplink FER as a function of uplink RSS is shown in Figure 5.13. As expected, the FER decreases as the RSS increases. The FER was below 1% for RSS values of -108 dBm and higher.

5.4 Handoff Testing

Handoff testing was performed by driving the measurement van on routes that would pass through the coverage area of each cell site and each sector within a cell site. Three routes were chosen for the handoff testing; two on Broadway, and one on Foothills Parkway. The first Broadway route was from 27th Way & Broadway north to US 36 (28th Street). The second Broadway route was from 27th Way and Broadway to Marshall Road. The third route went from 63rd Street and the Diagonal Highway south along the Foothills Parkway then onto US 36 to the Louisville exit. All routes were driven in northbound and southbound directions. Only the Ericsson mobile unit was used for handoff testing.

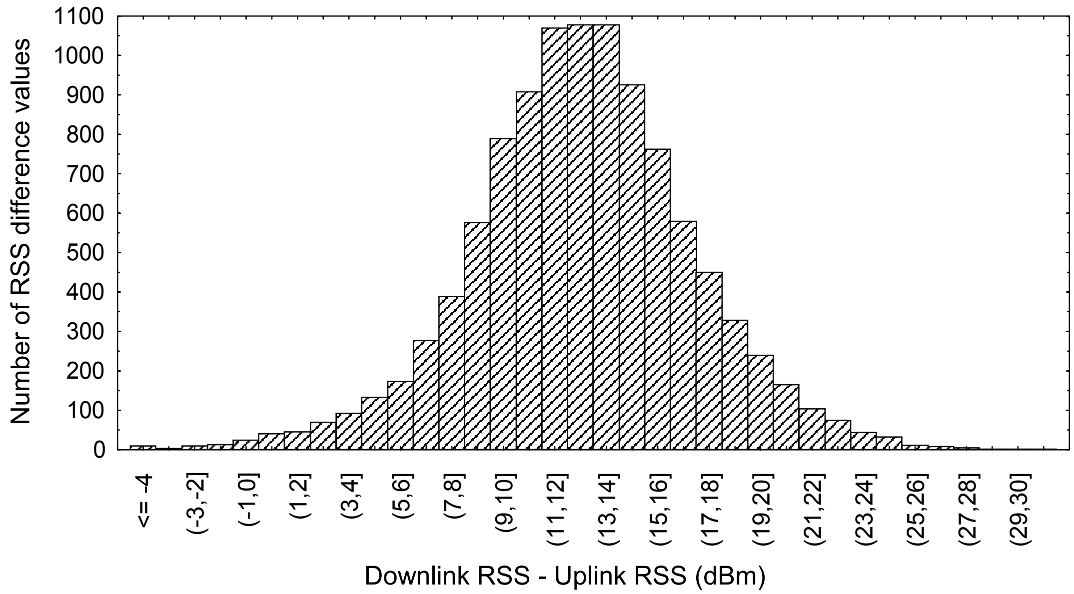

Figure 5.11. Histogram of difference between downlink and uplink received signal strength (RSS; TAG 4, WCO cell, Ericsson mobile unit).

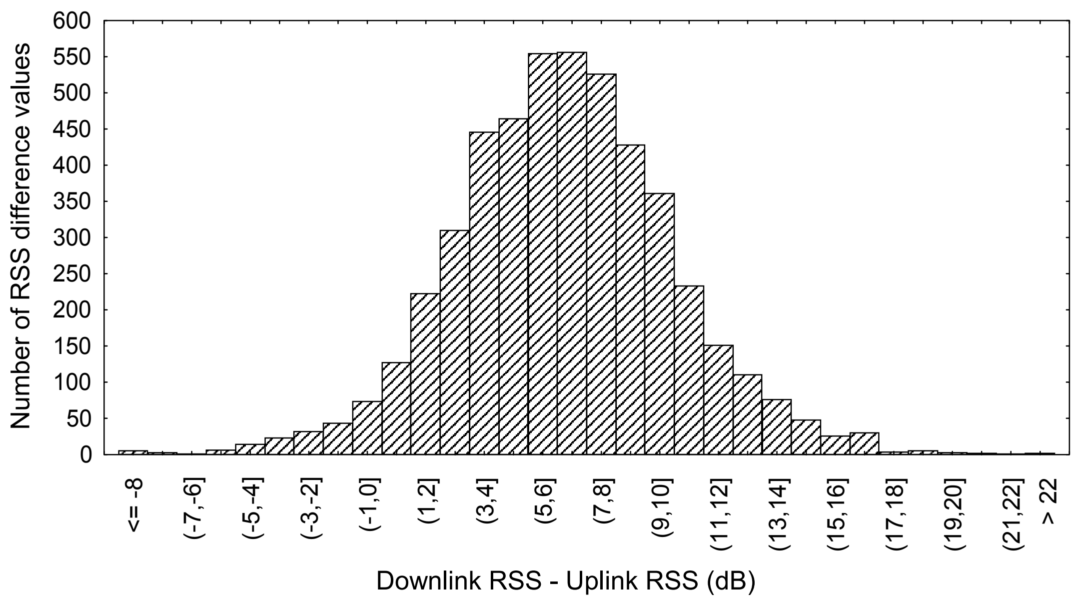

Figure 5.12. Histogram of difference between downlink and uplink received signal strength (RSS; TAG 4, WCO cell, Nokia mobile unit).

For the Broadway routes, the north sector of the TMCO cell, the north and southeast sectors of the GMM cell, and the southeast and southwest sectors of the WCO cell were active. For the Foothills Parkway route, the north and southwest sectors of the TMCO cell, the north and southeast sectors of the GMM cell, and the southeast sector of the WCO cell were active.

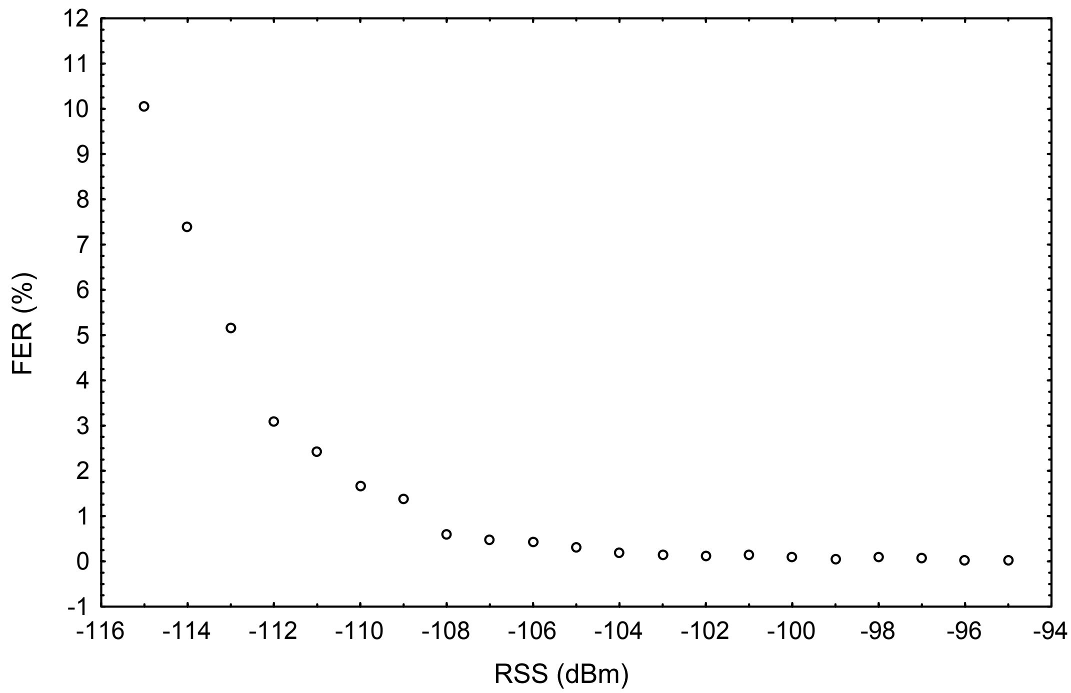

Figure 5.13. Uplink frame error rate (FER) vs. received signal strength (RSS; TAG 4, WCO cell).

Data collected from the handoff testing were analyzed to determine the change in the RSS values before and after handover. The time between successive handoffs was also examined. This gives an indication of the occurrence of any ping-ponging. Recall that ping-ponging was defined as a rapid switching of handoffs back and forth between cell sites and/or sectors. A total of 62 handoffs were examined in the analysis that follows.

Figure 5.14 shows the histogram of the change in RSS value before and after handoff. As in the handoff testing for TAG 5, some negative values of the change in RSS value before and after handoff occurred. Negative values represent cases when the average RSS value was actually lower after handoff than before handoff. In these cases, while the average RSS value in the current cell (or cell sector) was expected to be less than the average RSS value in the candidate cell (or cell sector) when handoff was initiated, by the time handoff was completed the average RSS in the new cell was less than the average RSS in the original cell.

Figure 5.15 shows the histogram of time between successive handoffs. The time between the handoffs is a function of speed, terrain configuration, route chosen, and cell planning. The occurrence of ping-ponging is indicated when the time between successive handoffs is too short (less than 10 s). Figure 5.15 shows that the time between successive handoffs was less than 10 s for 10 out of the 62 handoff cases. These were generally recorded on the uplink, on the Foothills Parkway route, directly east of the WCO.

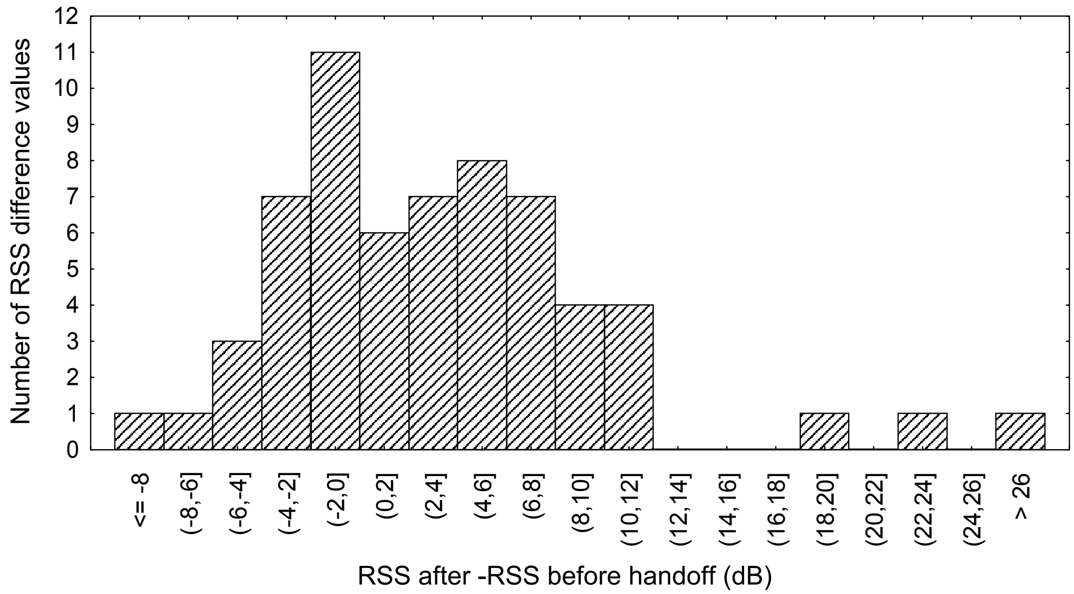

Figure 5.14. Histogram of change in received signal strength (RSS) during handoff (TAG 4).

5.5 Interference Testing

Both co-channel and adjacent channel interference measurements were performed for the downlink only. This provided C/I performance characterization for the mobile unit only. No interference measurements were taken for the base station receiver. The downlink measurements were made as the mobile unit traveled along routes that included areas with a range of good to poor C/I levels. The Ericsson mobile unit was used for all interference measurements.

For the co-channel interference measurement, the WCO north sector was used as the intended source. The GMM north sector was used as the source for the co-channel interference. Both the GMM north sector and the WCO north sector were tuned to the same frequency. A second carrier, at a different frequency, was transmitted from the GMM north sector. This carrier was used to measure the RSS of the interfering signal from the GMM site at the mobile unit. Note that in this configuration, the channel used for monitoring the interference level fades independently of the actual interfering signal.

For the adjacent channel interference measurement, the WCO north sector was again used as the intended source. The GMM north sector was used as the source for the adjacent channel interference. The carrier for the adjacent channel interferer was set 30 kHz away from the intended source.

Figures 5.16 and 5.17 show average BER class as a function of the C/I for both the co-channel and adjacent channel cases, respectively. Average BER class is computed by converting all the BER class values for a given C/I to BER values, averaging the BER values, and then finally converting the averaged BER value back to BER class. As expected for both the co-channel

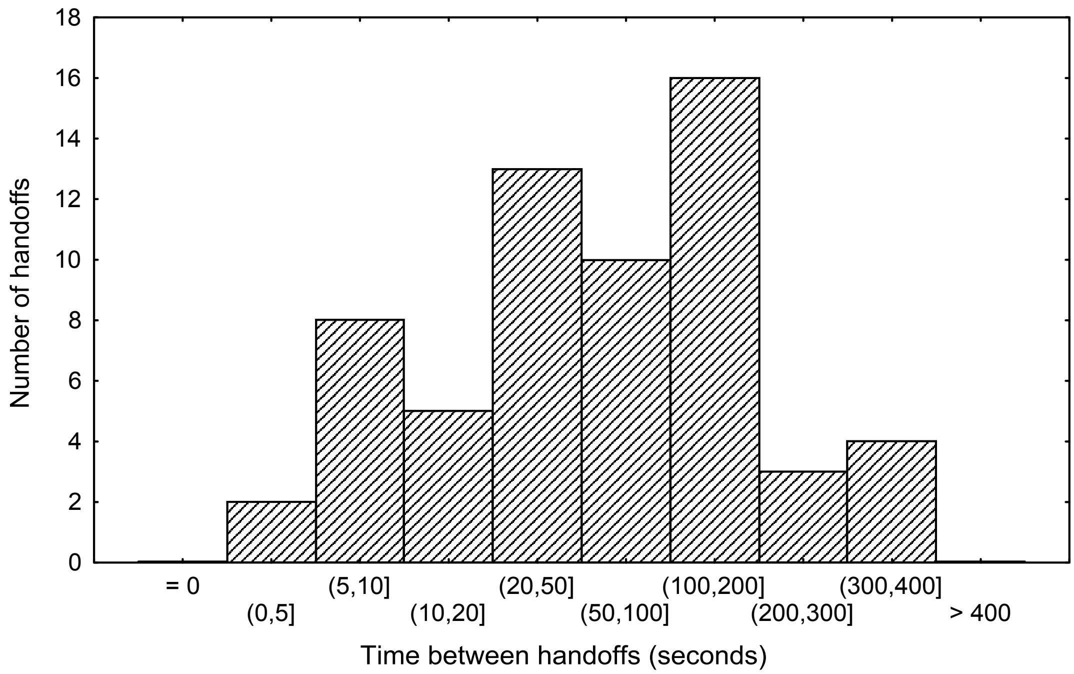

Figure 5.15. Histogram of time between successive handoffs (TAG 4).

and adjacent channel interference cases, when the C/I was increased, lower values of the BER class were seen, i.e., the BER decreases. Also as expected, to achieve the same BER class value, a much lower C/I was required for the adjacent channel interference than for the co-channel interference. As an example, for a BER class of 4, the co-channel C/I was 32 dB greater than the adjacent channel C/I.

5.6 Voice Quality

As discussed in Section 3.6, two types of voice quality measurements were made for the PCS JTC technology field trials in general: quasi-stationary measurements and handoff measurements. Both types of measurements were performed for the IS-136-based TDMA (TAG 4) technology.

5.6.1 Quasi-stationary Measurements

Voice recordings and various objective measures including uplink and downlink RSS, uplink number of bits in error per second, and downlink BER class were collected at locations on a 0.5-mi grid that encompassed the expected coverage area for the TMCO and WCO sites. Measurements were taken at 60 of the 82 locations that were identified for the quasi-stationary measurements as discussed in Section 3.6.1. The specific locations used are shown on the map in Figure 5.18. When making measurements at locations centered around the WCO cell site (the northern locations shown on the map in Figure 5.18), all sectors of the WCO cell were activated along with the north sector of the GMM cell site. When making measurements at locations centered around the TMCO cell site (the southern locations shown on the map in

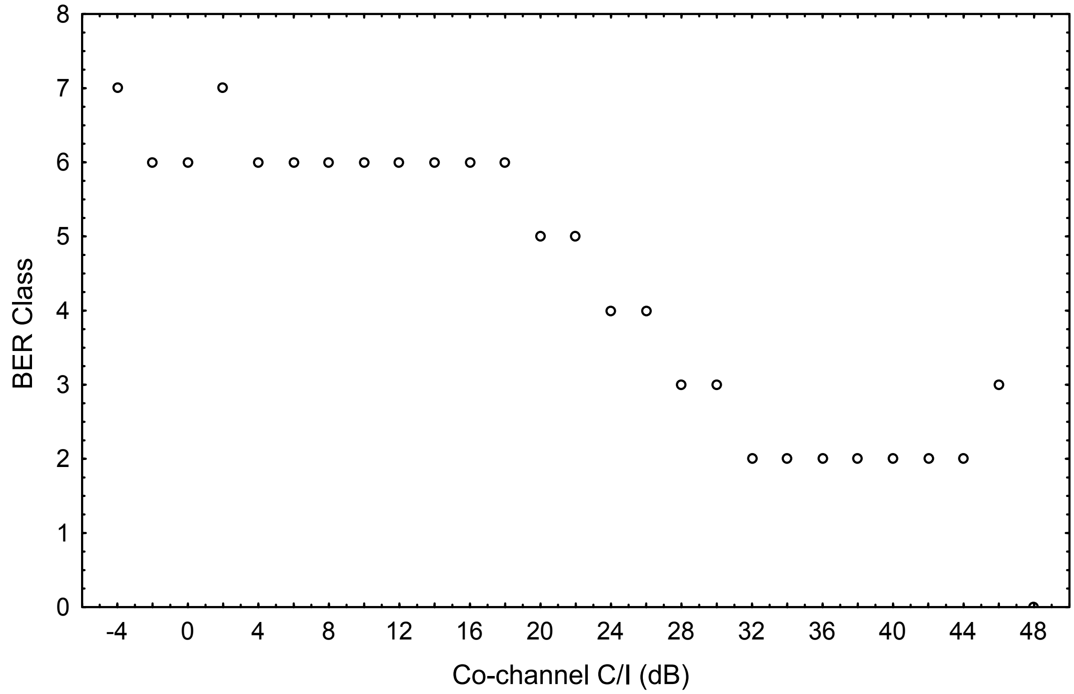

Figure 5.16. Average bit error rate (BER) class vs. carrier-to-interference ratio (C/I) for the co-channel interference case (TAG 4).

Figure 5.18), all sectors of the TMCO cell were activated along with the southeast sector of the GMM cell site.[9] At each location, data were collected as the measurement van traveled at one of two speeds. The vehicle traveled either 10 m or 100 m over the sample time. The particular vehicular speed used at each location (distance traveled over a fixed sample time) is shown on the map in Figure 5.18. At each location the sector that offered the highest signal strength was determined. This sector was used for the entire duration of the measurement at that location, i.e., handoff was not allowed during the measurement at a given location.

The measurements were taken at each location by establishing a call between the mobile and landline telephones. While the measurement van was in motion, an audio source tape was transmitted over the radio link in either the uplink or downlink direction. The source tape transmitted over each link was the same as that used for TAG 5 testing (see Section 3.6.1).

The received voice transmissions were recorded on digital audio tape at the receiver for the uplink and at the receiver for the downlink. (Note that these recordings were not taken simultaneously.) The recorded voice segments were then digitized with 16-bit resolution at a 22-Ksample/s rate and stored on a hard disk drive. At each location, the objective measures including uplink and downlink RSS, uplink number of bits in error per second, and downlink BER class were also collected. The objective measures were collected at the same time as the voice recordings.

[9] The activation of cells and cell sectors for the quasi-stationary measurements in TAG 4 testing was slightly different from the testing in the previous JTC PCS field trials. In the TAG 2 testing, both the TMCO and WCO cells were activated. In the TAG 5 testing, only one cell was activated at a time.

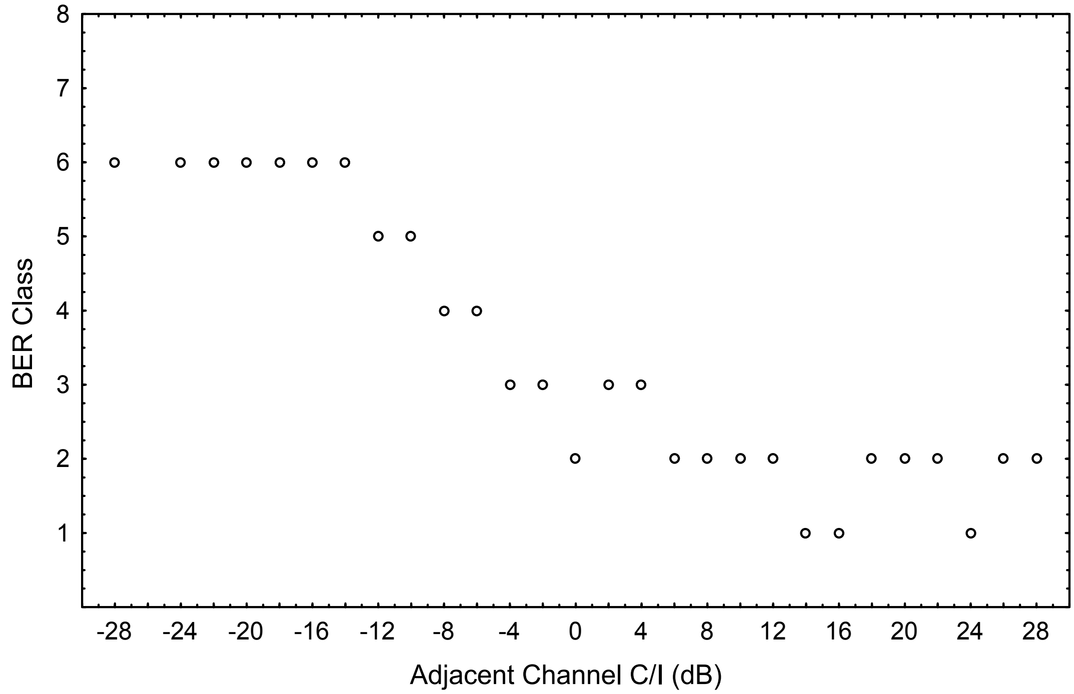

Figure 5.17. Average bit error rate (BER) class vs. carrier-to-interference ratio (C/I) for the adjacent channel interference case (TAG 4).

For the quasi-stationary measurements, voice quality of the voice segments was determined by both mean opinion score (MOS) and expert listener techniques. The following sections discuss these techniques and present the results based on the application of these techniques.

5.6.2 Mean Opinion Score Assessment

To accomplish the MOS testing, a pool of 32 subjects was recruited from the Boulder, Colorado area. Each of the following age groups were represented within the subject pool: 18-25, 25-35, 35-45, 45-55, and those over 55 years of age. There were an equal number of male and female subjects. The subjects were cordless, noncellular telephone users.

Four groups consisting of eight subjects each from the subject pool were formed The subjects were asked to rate voice segments by answering the three questions listed in Section 3.6.2 after each segment was presented.

First, the subjects in each of the four groups were presented 10 practice voice segments to rate. The practice segments included two 64-kbps wireline voice segments (actually one 64-kbps wireline segment presented twice) and eight voice segments of varying degrees of quality collected from field measurements. It was also important to allow listeners to be exposed to the types of distortions inherent in the TAG 4 voice coding performance at varying levels. The goal was to demonstrate a wide range of quality in the practice segments presented

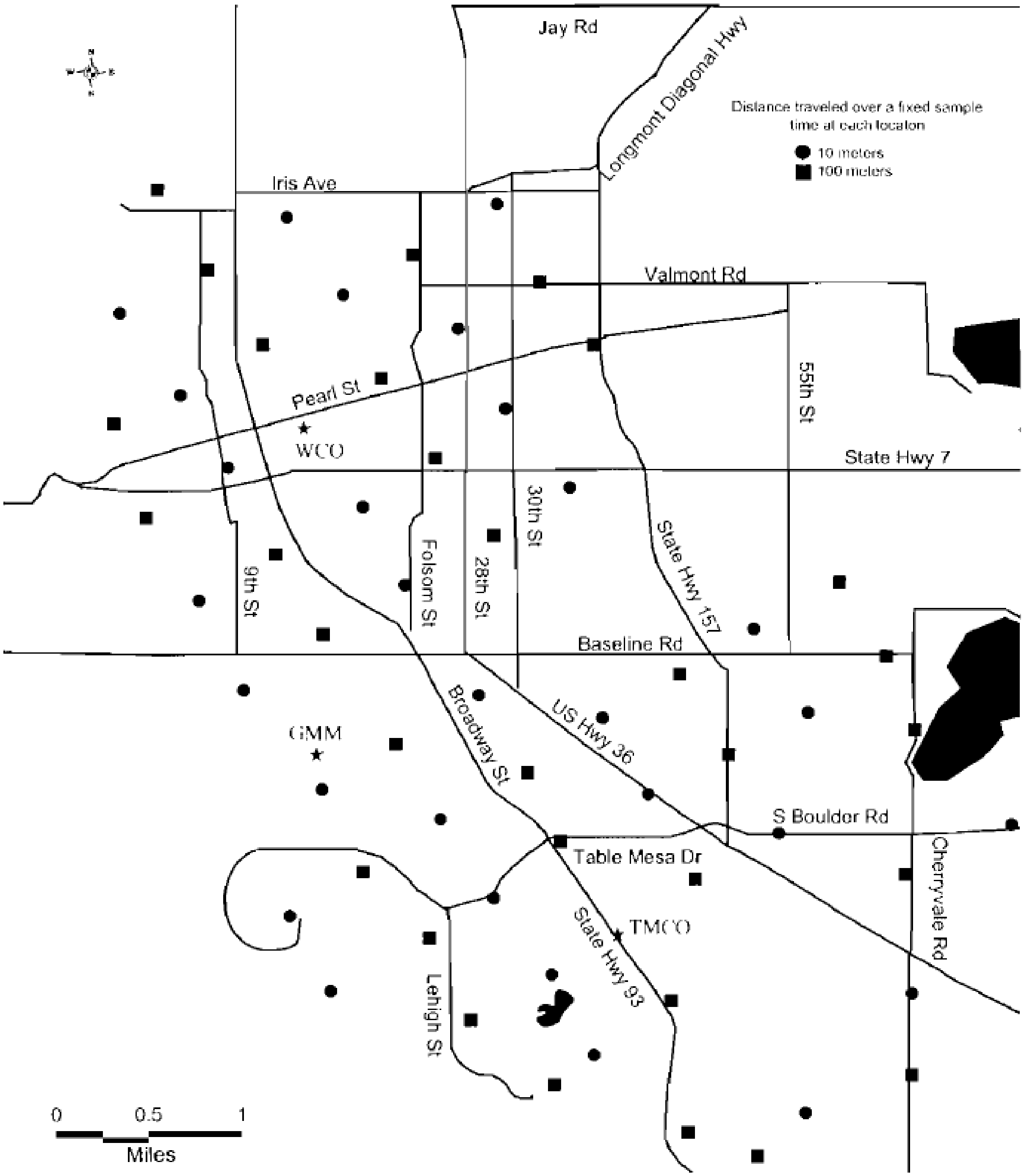

Figure 5.18. Quasi-stationary measurement locations and vehicular speed used at each location for TAG 4.

to listeners. It was felt that the voice segments collected during the quasi-stationary measurements did not represent the full range of quality needed, therefore during field testing, extra voice segments were collected while adjusting the transmit power within a cell. The practice segments allowed inexperienced listeners to gain exposure to the full range of possible voice quality, from excellent to poor, and ensured proper scaling of the MOS's.

After the practice segments were presented, the subjects in each of the four groups were asked to rate 33 voice segments: 3 reference voice segments and 30 voice segments from the field trial measurements. Every listener within a group listened to the same voice segments. The reference voice segments were either 64-kbps wireline segments or poor quality voice segments collected from the field. The 30 voice segments from the field trial measurements came from the uplink or downlink measurements at a portion of the 60 measurement locations.

A total of 120 voice segments were collected from the field measurements, representing the uplink and downlink measurements at all 60 measurement locations. Out of the 120 voice segments, nine of the voice segments recorded during the downlink measurements suffered from an unrealistically poor voice quality (possibly due to electrical interference being coupled in at the input to the digital audio tape recorder). Therefore, these voice segments were not included in the data analysis that follows. The number and type (male or female) of sentences that were used and the order in which they were used to form a voice segment are the same as that described in Section 4.6.2. Subjects were given two 10-min breaks half way through the session. After all segments were presented, subjects filled out a post-trial questionnaire.

For each voice segment, voice quality ratings (answers to the question "How would you rate the overall quality of the sound?") from each subject within a group were averaged to obtain an MOS. The results from all four of the groups, from a total of 111 voice segments (because 9 out of the 120 voice segments were not used as explained earlier), are shown in the histogram in Figure 5.19. Overall, the voice segments were marked favorably, with 90% of the segments rated between fair and excellent. The average MOS was 3.53 and the standard deviation was 0.49. The one very low MOS occurred at a location where no LOS propagation existed between the mobile unit and the base station at either the TMCO or GMM cell sites. This location is in a small valley behind both cell sites.

Figures 5.20 and 5.21, show histograms of MOS's for the uplink and downlink, respectively. The data shown in these histograms are from the uplink and downlink voice segments at 51 out of the 60 locations. Both the uplink and downlink voice segments from the remaining nine locations were not used in these histograms. The average MOS for the uplink was 3.41 and for the downlink was 3.68. A t-test revealed that there is a statistically significant difference in the average MOS's between the uplink and downlink. The downlink MOS's were noticeably better than the uplink MOS's as expected because of the large link imbalance as discussed in Sections 5.3.1 and 5.3.2.

The relationship between the MOS's and some of the objective measures was initially investigated by generating some scatter plots. Figure 5.22 shows the relationship between the MOS's and average RSS for both the uplink and downlink combined (111 voice segments). In Figure 5.22, some variation in MOS's is seen for all values of average RSS. Figures 5.23 and

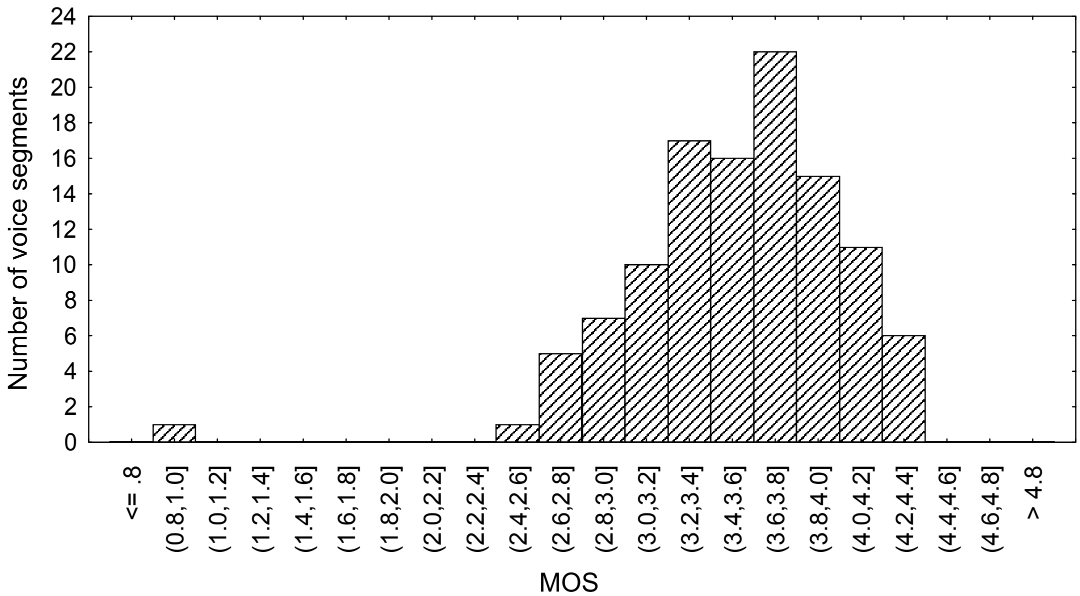

Figure 5.19. Histogram of mean opinion scores (MOS's) for all voice segments (TAG 4).

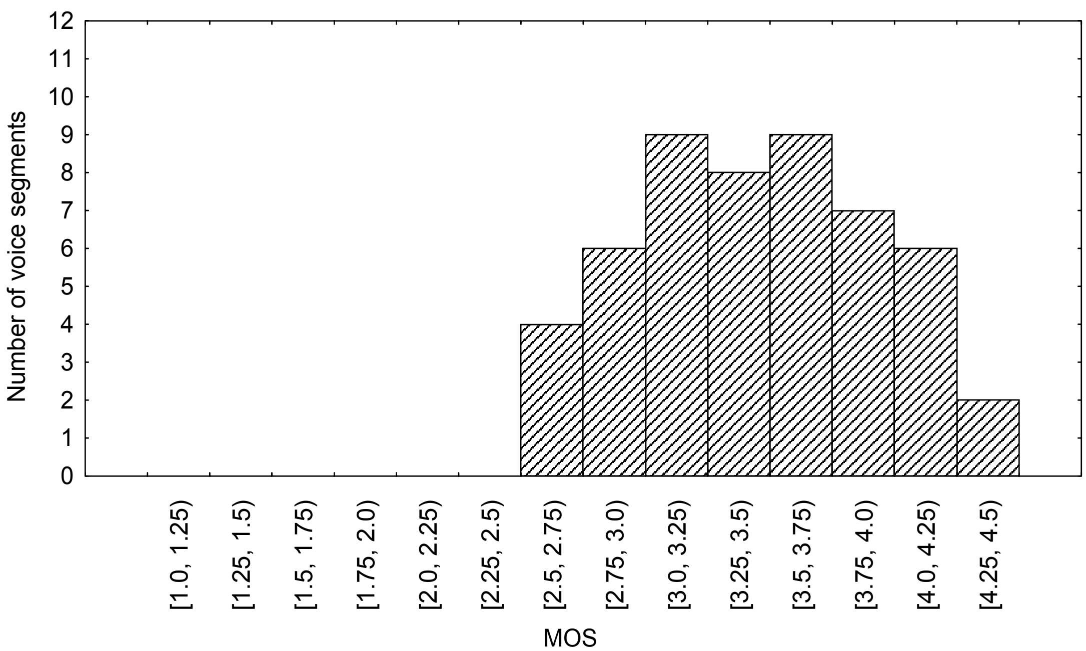

Figure 5.20. Histogram of mean opinion scores (MOS's) for the uplink (TAG 4).

5.24 show the relationship between the MOS's and the average BER for both the uplink (all 60 voice segments) and downlink (51 voice segments), respectively. Note that only one voice segment on the uplink had an average BER greater than 0.3%. The average BER never exceeded 0.4% on the downlink. While both the uplink and downlink show some variation in MOS's for all values of average BER, more variation is seen on the uplink. The variation in

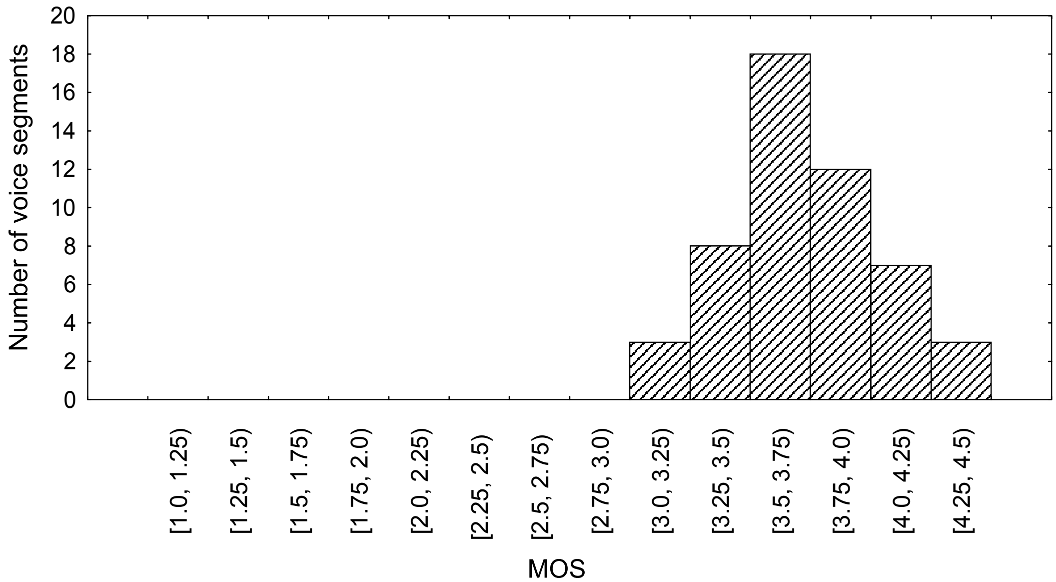

Figure 5.21. Histogram of mean opinion scores (MOS's) for the downlink (TAG 4).

MOS's for voice segments with average BER's on the uplink between 0 and 0.3% ranged from 2.50-4.38. The variation in MOS's for voice segments with average BER's on the downlink between 0 and 0.4% was relatively constant; MOS's ranged from 3.00-4.25.

Pearson product-moment correlations were performed to determine the correlation between MOS's and average RSS and MOS's and average BER (for the uplink and downlink data combined). The correlation coefficient between MOS and average RSS was 0.30 and that between MOS and average BER was -0.50. These correlation coefficients between MOS and averaged objective measures suggest that a strong linear relationship between MOS and the objective measures does not exist. The correlation between the average RSS and the average BER was -0.33. A higher correlation between these measures was expected.

While there does not appear to be a strong linear relationship between MOS and the objective measures, there still may be a consistently increasing or decreasing relationship between them. The Spearman rank correlation can be used to determine if a consistently increasing or decreasing trend may exist between MOS and the objective measures. Spearman rank correlations were performed to determine the correlation between the ranks of MOS and the ranks of average RSS and between the ranks of MOS and the ranks of average BER. The Spearman rank correlation coefficient (for the uplink and downlink data combined) between MOS and average RSS was 0.16 and that between MOS and average BER was -0.05. These very low rank correlations between MOS and averaged objective measures suggest that a consistently increasing or decreasing relationship between MOS and the objective measures does not exist.

Note that as in the TAG 5 and TAG 2 data analysis, the objective measures were averaged over the entire length of the voice segment. By analyzing the instantaneous variation or possibly minimum and maximum values of the objective measures within the voice segment, further insight may be gained on the behavior of MOS's.

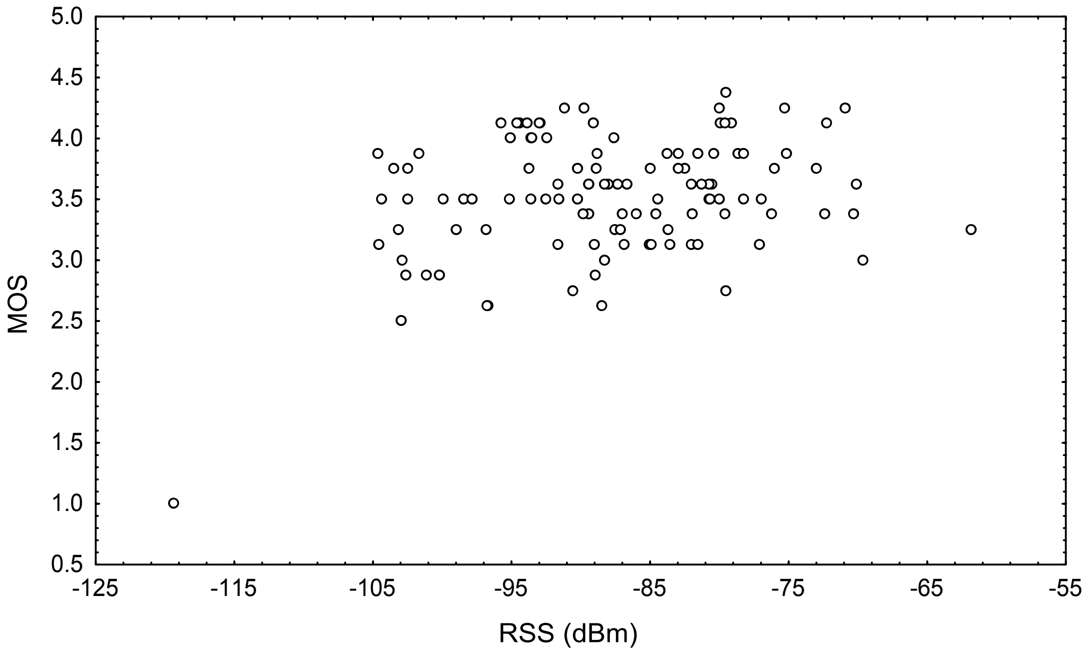

Figure 5.22. Mean opinion score (MOS) vs. average received signal strength (RSS; TAG 4).

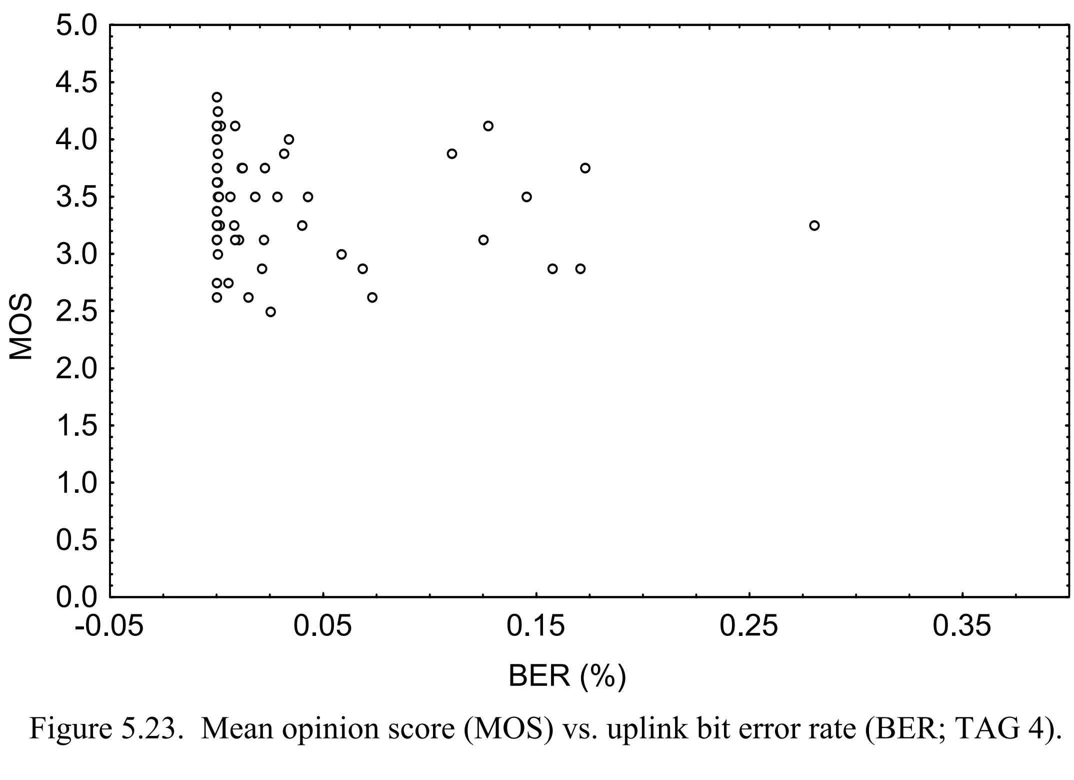

Figure 5.23. Mean opinion score (MOS) vs. uplink bit error rate (BER; TAG 4).

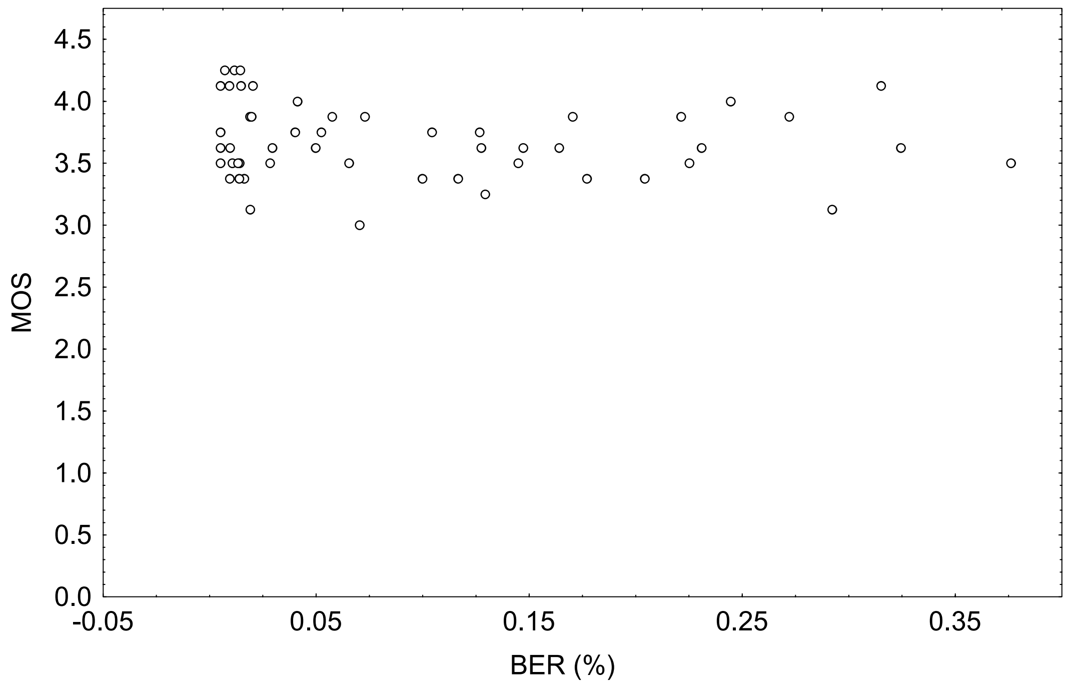

Figure 5.24. Mean opinion score (MOS) vs. downlink bit error rate (BER; TAG 4).

By gathering listeners' comments from post-test questionnaires, more information about the nature of MOS's was obtained. Namely, it is evident from questionnaires that there are several types of distortions in quality possible in the voice segments according to listeners:

1) echo;
2) "chirping," or "squeaking" noise; and
3) muting.

The nature of these distortions are likely judged differently by different listeners. Intelligibility and speaker recognition are two main aspects of perceived quality. In the case of most of the voice segments, intelligibility remained high. As a result, overall MOS's were high.

5.6.3 Expert Listener Assessment

In addition to being rated by listener panels in MOS testing, the voice segments were rated by an expert listener. The expert listener ratings followed the identical procedure as in the PCS 1900 (TAG 5) testing. This procedure is described in Section 3.6.3.

Figure 5.25 shows the relationship between expert listener ratings and percent acceptability (the percentage of listeners rating a given voice segment as acceptable). The boxes represent the middle half of the data (from the 25th percentile to the 75th percentile). The solid circles represent the median percent acceptabilities for each of the expert listener ratings. The lines extending out of the boxes depict the spread of the data.

For the most part, the definitely acceptable ratings of the expert listener were relatively good indicators of percent acceptability. In the definitely acceptable category, most of the voice segments were rated as acceptable by 70-100% of the subjects; a few voice segments were rated as acceptable by less than 70% of the subjects. The marginally acceptable and unacceptable expert listener categories show more variance. In the marginally acceptable category, few voice segments were rated as acceptable by 30-70% of the subjects. Similarly, in the unacceptable category, few voice segments were rated as acceptable by less than 30% of the subjects.

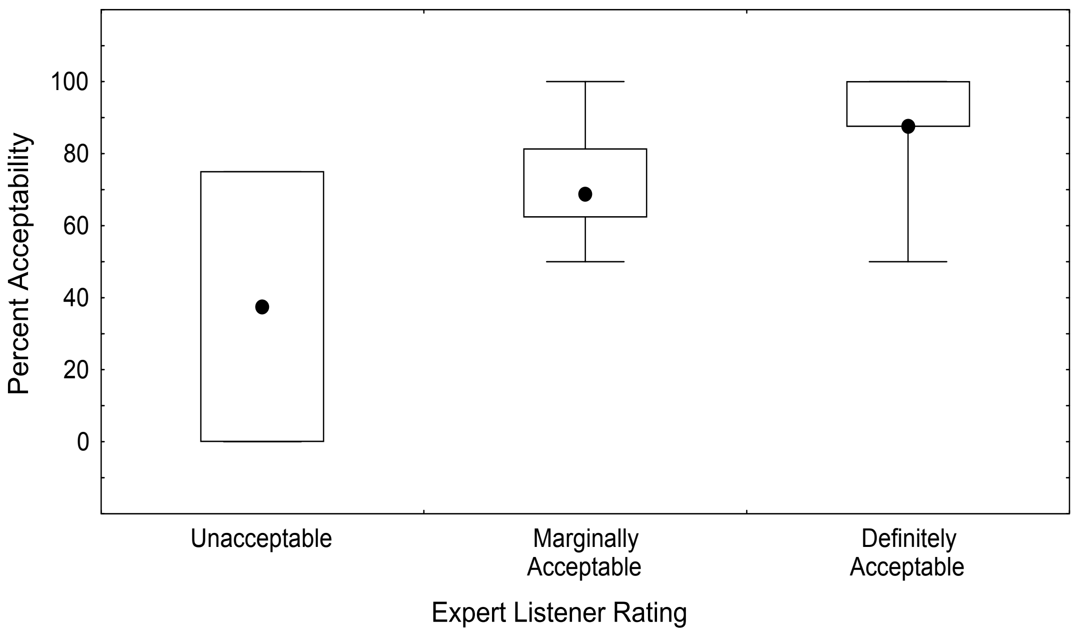

Figure 5.25. Percent acceptability vs. expert listener rating (TAG 4).

The Pearson product-moment correlation coefficient between MOS and percent acceptability was 0.80, showing a strong correlation between these measures, as would be expected. The Pearson product-moment correlation coefficient between MOS and expert listener rating was 0.53 and that between percent acceptability and expert listener rating was 0.53, indicating some correlation between these measures. However, a strong linear relationship between these measures does not exist.

It is possible that expert listener rating, average RSS, and average BER can be predictors of MOS when all are taken together. A multiple regression analysis was completed in order to determine if MOS is related to a combination of all the above factors. Interactions among the various factors were not considered in the analysis.

The result showed that only 36% of the variance can be explained by expert listener rating, average RSS, and average BER. The result of the multiple regression analysis also showed that all these factors reliably account for the 36% of variance. From this analysis, it appears that expert listener rating, average RSS, and average BER taken together do not offer any substantial predictive power over MOS's.

5.6.4 Voice Quality Handoff Measurements

Continuous voice recordings of transmitted voice signals were made as the mobile unit traveled along routes through handoff areas. While the measurement van was in motion, an audio source tape was transmitted over either the uplink or downlink. (Uplink and downlink voice recordings were made during separate runs.) The source tape consisted of Harvard sentences and was played continuously as the measurement van traveled along each route. The received voice transmissions were recorded on digital audio tape at the receiver for either the uplink or downlink. Continuous voice recordings were made along the route until the call was finally dropped. The routes were selected according to previous tests at the BITB. One route was driven along Broadway, at a vehicular speed of 25-30 mph from US 36 south to 27th Way. The second route was driven along Broadway from 27th Way south to Chambers Road or in some cases further south to Marshall Road. A third route was driven along Foothills Parkway, at a vehicular speed of 55-60 mph. The beginning of a route was selected within the coverage area. Four runs were conducted along each route: an uplink run traveling southward, an uplink run traveling northward, a downlink run traveling southward, and a downlink run traveling northward. Only the Ericsson mobile unit was used for the voice quality handoff testing. Voice quality was assessed using the expert listener methodology described in Section 3.6.3.

For the voice quality handoff testing, an expert listener rating was made for each 4-s period of the continuous voice recordings taken along a measurement route. In general, voice quality was good throughout most of the handoff testing. Voice quality did appear to degrade briefly during handoffs; the expert listener was able to identify when a handoff occurred. During handoff, a 200- to 400-ms period of muting was evident.

5.7. Manufacturers' Statements

Statements provided by the manufacturers involved in the testing are included in this section. These statements are identical to those given in [4], except for some minor editorial changes.

5.7.1 AT&T Network Wireless Systems

AT&T Network Wireless Systems would like to thank U S West, NTIA, Ericsson, Nokia, and McCaw (now AT&T Wireless) for their hard work and support during the testing of PCS TDMA. Special thanks is extended to U S West and NTIA for their contributions during the field trial, and for providing the BITB. AT&T Network Wireless Systems would also like to thank Nokia and Ericsson for providing mobile equipment to demonstrate the performance of PCS TDMA. Individuals from all of the above organizations put in many long hours cooperating together to meet the tight test schedules.

AT&T was pleased to supply the wireless infrastructure to demonstrate the PCS TDMA air interface as part of the JTC standardization process for PCS band air interfaces.

As can be seen from this section of the report and the Visitors' Day presentations and demonstrations, the PCS TDMA system performed beyond expectations. The PCS TDMA field tests consistently demonstrated high voice quality and excellent coverage.

Due to the nature of this technology demonstration, and the time constraints in the planning and execution of this event, external upbanders were utilized both on the base station infrastructure, and the mobile equipment. Time constraints, as well as the nature of the testing resulted in minimal time to tune the system after installation, and run proper interoperability, calibration, and integration tests prior to our arrival in Boulder. As a result, the links were not balanced, power control was not implemented and optimized, and the system was not tuned optimally as would have occurred had this been production equipment in a service provider environment.

Due to the JTC's legitimate need to standardize the testing for all of the PCS technologies as much as possible, cell site selection was predetermined by the BITB, and was not at TAG 4's discretion. As a result, the cell site selection was not optimized to cover the test area for TDMA.

As with all high technology, new products and ideas such as PCS TDMA production minicells and improved speech codecs are in development and hitting the marketplace as this is written.

5.7.2 Ericsson, Inc.

Ericsson Incorporated would like to thank U S West and ITS for their support during the demonstration of the performance of the US TDMA PCS radio air interfaces at 1.9-GHz at the BITB in Colorado. In particular, the support of M. Laflin of ITS and the technical staff of U S West is acknowledged and very much appreciated. Ericsson would further like to thank all of the AT&T Wireless Network Systems participants for their hard work in completing the compatibility requirements for the 1.9 GHz TDMA PCS system. In particular we wish to thank J. Siskind for his hard work in support of the system test.

Ericsson was pleased to provide equipment to demonstrate the TDMA PCS system in the BITB, and we feel that the TDMA system performed quite well as the report shows. The results were in agreement with the previous TDMA field demonstrations in other locations.

Because of the site location constraints and the time constraints, there was no attempt to optimize the system used in Boulder; however, the TDMA test system provided consistently high voice quality.

The BITB test system mobile units provided consisted of noncommercial engineering equipment based on the IS-54B standard. Work is underway to provide hand-held terminals for the commercial deployment of the TDMA PCS standard.

5.7.3 Nokia Mobile Phones

Nokia Mobile Phones, Inc. would like to thank AT&T Wireless Network Systems, U S West, and ITS for their support during the demonstration of the performance of the 1900-MHz US TDMA radio air interface at the BITB in Colorado. We found the atmosphere and attitude during the tests extremely positive and co-operative.

The cell sites were not completely tuned and balanced; that would have required more time than was allowed by the schedule. However, the purpose of this test was to prove the basic system functionality and to define the coverage area, and this was successfully done.

Noncommercial equipment was used for the testing. However, Nokia Mobile Phones is doing development work to provide handportable phones for IS-136-based PCS systems.

This test proves that the IS-136-based PCS system is a very competitive solution for the PCS operators, especially when the new vocoder becomes available for the system.

6. TAG 7 (WIDEBAND CDMA) TESTING

This section describes the test plan, methodology, and results for the technology field trial conducted by TAG 7. The technology tested was the Wideband CDMA PCS standard. The base station and mobile equipment were provided by OKI America and Berkeley Varitronics Systems, Inc. The TAG 7 field tests examined area coverage and voice quality under quasi-stationary conditions. No handoff testing or interference testing was performed.

The information presented in this section is taken from [5]. The reader is referred to [5] for a more complete and detailed presentation of the TAG 7 technology field testing at the BITB.

6.1 TAG 7 Test System Configuration

The test system for TAG 7 was different from the previously tested JTC PCS systems (TAG 5, TAG 2, and TAG 4) because it consisted of a single base station and a single mobile unit. The base station was set up to operate on one sector of one cell at a time while testing proceeded for that particular sector. All three sectors of both the TMCO and WCO cell sites were tested one at a time. Because only one base station was used for the TAG 7 testing, no network configuration to connect base stations was needed. The voice signal interface was provided directly into the base station without access to any network switching or transmission components. No antenna diversity was used for any of the TAG 7 testing.

6.2 Calibration

A calibration of the base station RSS was performed by injecting a digitally modulated signal of known level into the receiver and then generating a table of scale factors that were used by the logging software to provide accurate RSS reports during the field trial.

The input signal into the base station receiver was provided by the mobile unit through a coaxial cable and high-power and variable attenuators. Losses through the coaxial cable and attenuators were measured and a correction, or offset, was added to the signal power output by the mobile unit. The level of the injected signal was varied over the dynamic range of the base station receiver to generate the table of scale factors. The maximum difference between the actual RSS and the RSS reported at the base station receiver was found to be 1.5 dB.

The OKI mobile units that were used in the TAG 7 field trial were checked for accuracy of the RSS values reported. This was done by injecting a digitally modulated signal of known level directly into the antenna port of the mobile unit and comparing this signal level with the RSS value reported by the mobile unit. The input signal into the mobile unit receiver was provided by the base station through a coaxial cable and high-power and variable attenuators. Losses through the coaxial cable and attenuators were measured and a correction, or offset, was added to the signal power output by the base station. The level of the injected signal was varied over the dynamic range of the mobile unit's receiver to find the maximum difference between the actual RSS and the RSS reported by the mobile unit's receiver. The maximum difference was found to be 2.5 dB. (Note that the OKI mobile units report RSS in 0.5 dB steps.)

Other parameters such as base station transmitter characteristics were not examined during this calibration.

6.3 Area Coverage Testing

The base station and mobile unit power were set so that the output power was approximately 5.6 W and 0.25 W ERP, respectively. Measurements to show area coverage were taken with the mobile unit located in a mini-van as in the previous JTC PCS field technology trials (TAG 5, TAG 2, and TAG 4). The mobile unit was mounted inside the van on the same wooden structure used in all of the JTC PCS technology field trials. This structure is described in Section 3.3. The measurements were taken by driving along routes (radials) away from the cell site. When time permitted, measurements along additional routes in between the radials were taken.

Since only one base station and mobile unit were used in the testing, only one sector of one cell site was activated at a time during area coverage testing. The data were collected both at the mobile unit and at the base station. The data collected at the mobile unit included GPS location, velocity, and time; downlink RSS; downlink BER; and other system parameters. The data collected at the base station included time, uplink RSS, uplink BER, and other system parameters.

Calls were originated prior to the start of data collection. Collection of mobile unit data and base station data was initiated as the measurement van began traveling away from the cell site along the drive route. The vehicle speed was approximately 5 mph. At the end of the route, the data collection was stopped and the data were saved to disk. All data collected were averaged over one second. The mobile unit was capable of open loop power control; however, the open loop power control was not used in this testing.

6.3.1 TMCO Area Coverage Data

The test procedure followed and data collection methodology used are explained in Section 6.3. Downlink RSS as a function of distance is shown in Figure 6.1. The large variation in RSS seen in Figure 6.1 is mostly due to shadowing. Note that signals up to approximately -87 dBm exist out to approximately 0.83 mi from the cell site.

A rough estimate of the coverage area was determined by assuming that an RSS of -83 dBm or greater is desired. The measured RSS data along all of the routes driven within the cell were used to determine the coverage area. Due to the irregularity of the terrain, the RSS varied significantly along the TMCO routes, crossing the -83-dBm level several times before finally staying below -83 dBm. The point along each route where the RSS first dropped below -83 dBm was used to define the coverage boundaries. For this case, the coverage boundaries were approximately 0.55 mi due north, 0.3 mi due southeast and southwest, and 0.6 mi due south.

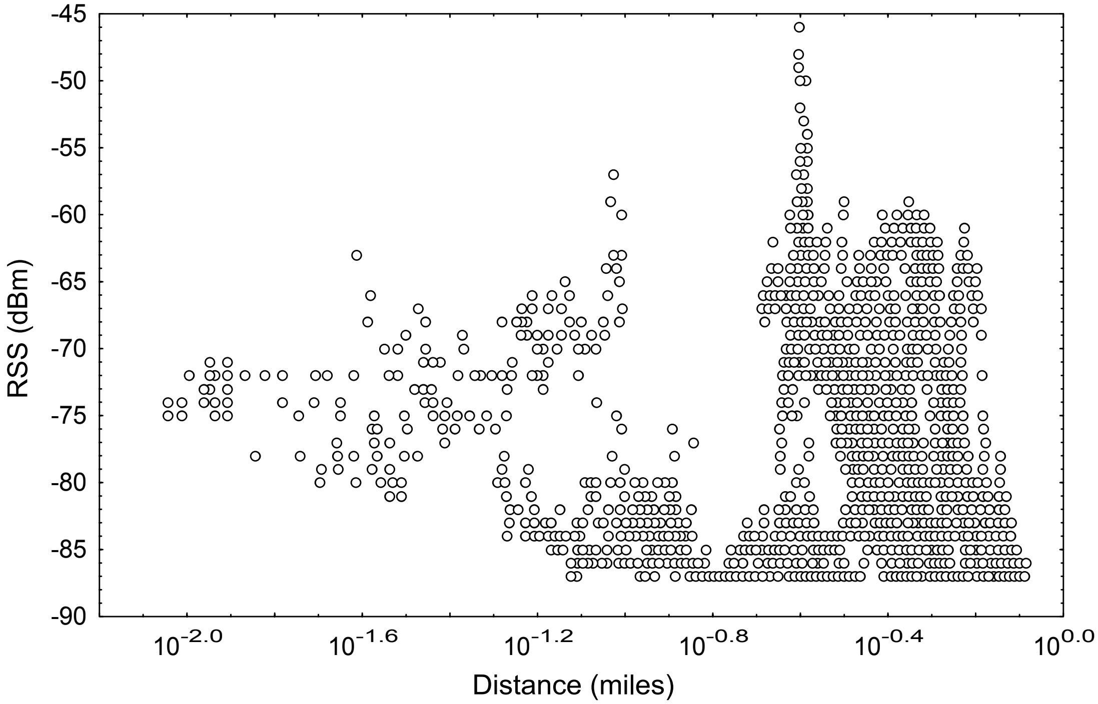

Figure 6.1. Downlink received signal strength (RSS) vs. distance (TAG 7, TMCO cell).

Figures 6.2 and 6.3 show the histograms of uplink and downlink RSS values, respectively. The direct comparison between uplink and downlink (link balance) was not possible, due to misalignment of the time stamps in the uplink and downlink data files. However, as seen in Figures 6.2 and 6.3, the histograms of RSS data are similar for both links. For both links, the majority of RSS values were less than -80 dBm. Also, for both links, RSS values greater than -80 dBm were fairly evenly distributed. The lowest recorded RSS was -87 dBm for the downlink and -85 dBm for the uplink.

Figures 6.4 and 6.5 show the uplink and downlink BER histograms, respectively. For the uplink, the vast majority of BER values were less than or equal to 0.001%. For the downlink, while a large number of BER values were less than or equal to 0.001%, there were more BER values in the 0.01 - 0.1% and 0.1 - 1.0% ranges than for the uplink.

6.3.2 WCO Area Coverage Data

As for the TMCO cell, the test procedure followed and data collection methodology used are explained in Section 6.3.

Downlink RSS as a function of distance is shown in Figure 6.6. As seen in the TMCO cell, there is a large variation in RSS for a given distance. This is due to shadowing and the choice of routes. Note that signals up to approximately -87 dBm exist approximately 0.89 mi from the cell site.

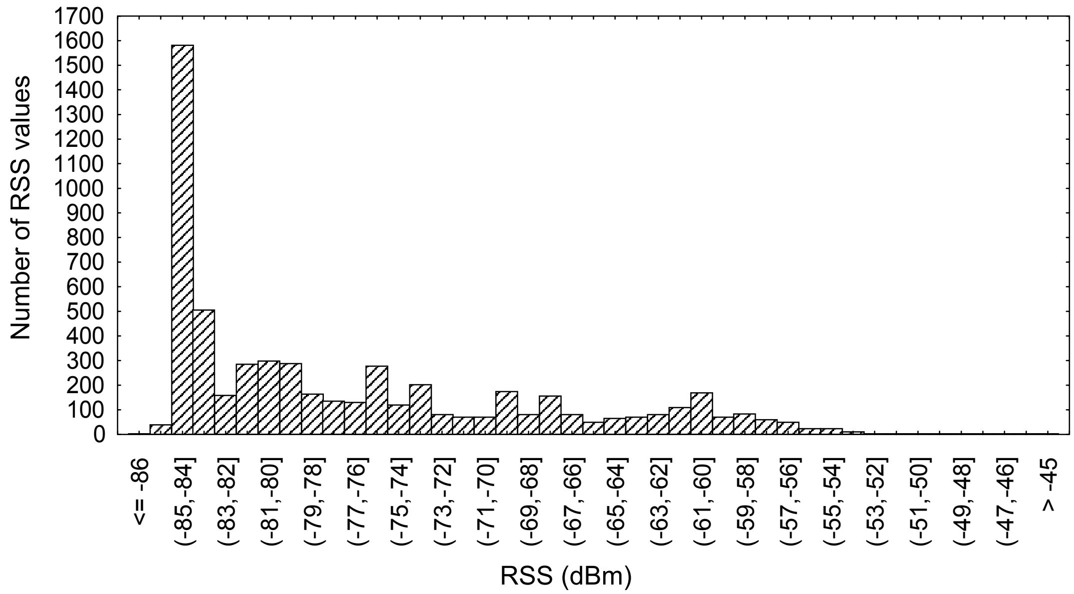

Figure 6.2. Histogram of uplink received signal strength (RSS) values (TAG 7, TMCO cell).

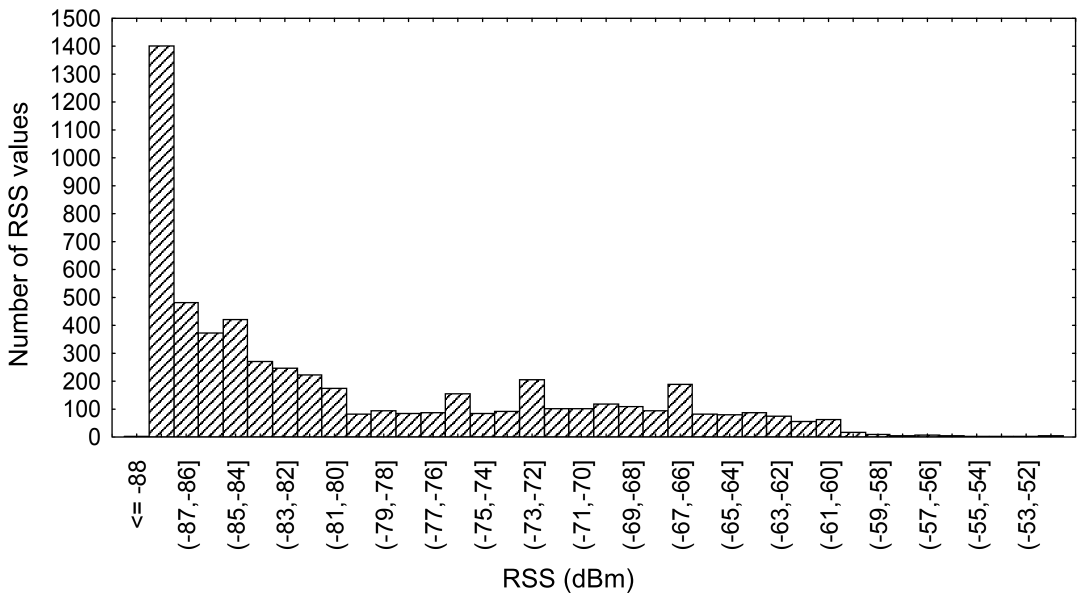

Figure 6.3. Histogram of downlink received signal strength (RSS) values (TAG 7, TMCO cell).

The coverage boundaries were determined in the same manner as for the TMCO cell. The coverage boundaries for the WCO cell were approximately 0.36 mi due north, 0.5 mi due east, and 0.42 mi due south. The WCO cell had a smaller total coverage area than the TMCO cell. The differences in area coverage between the two cells are due to different types of environments and different terrain profiles.

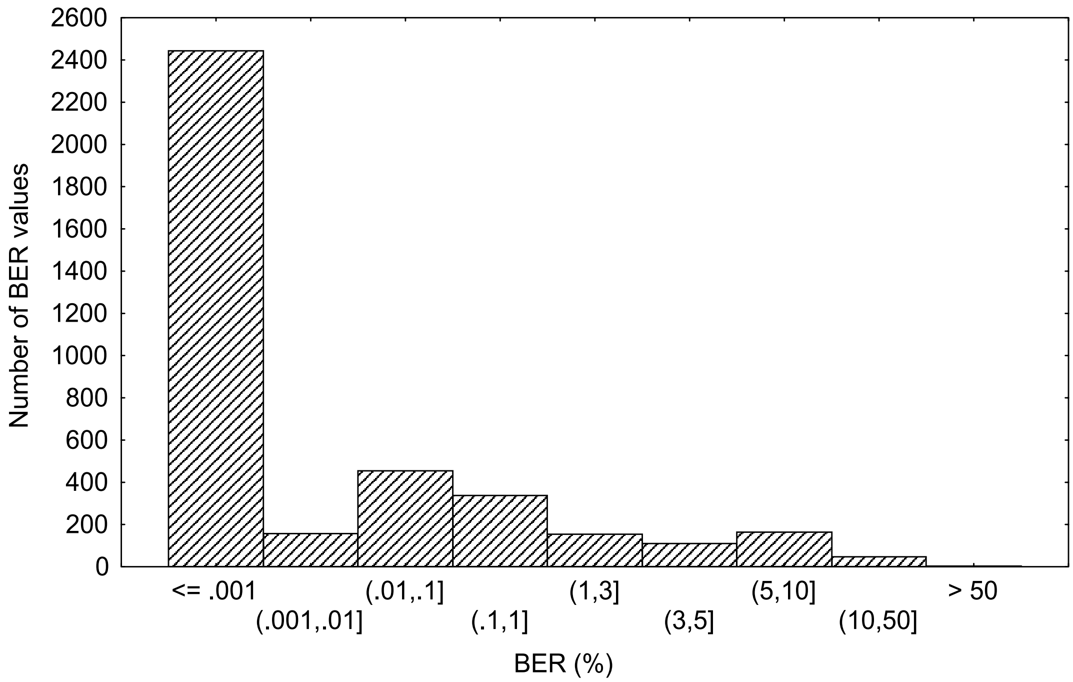

Figure 6.4. Histogram of uplink bit error rate (BER; TAG 7, TMCO cell).

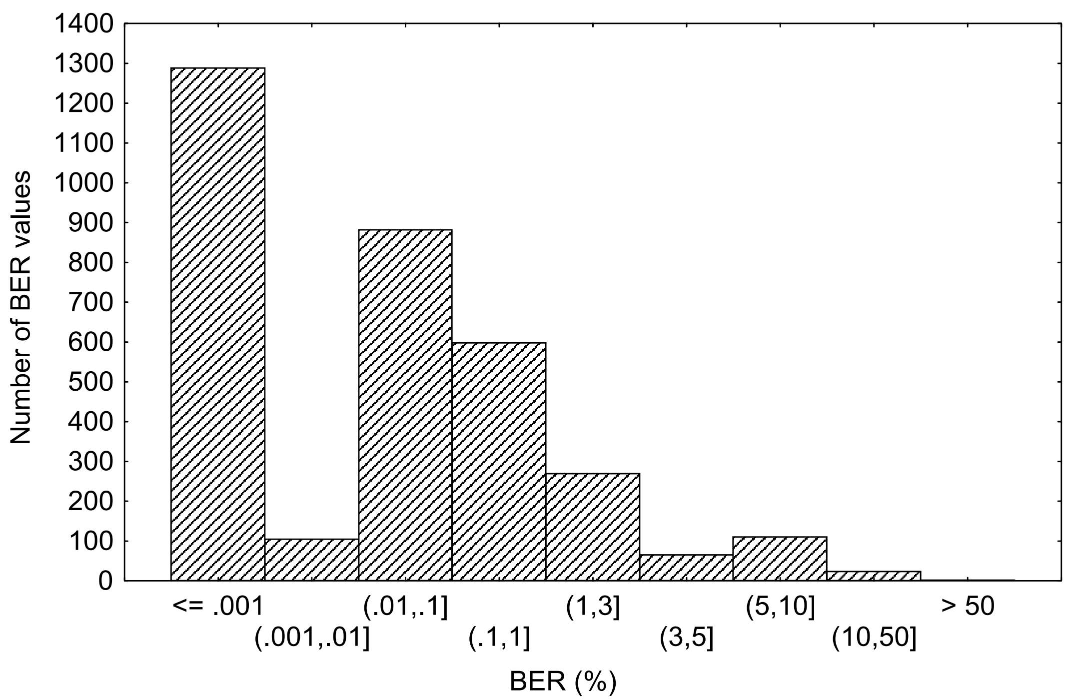

Figure 6.5. Histogram of downlink bit error rate (BER; TAG 7, TMCO cell).

Figures 6.7 and 6.8 show the histograms of uplink and downlink RSS values, respectively. As in the TMCO data, the direct comparison between uplink and downlink (link balance) was not possible, due to misalignment of the time stamps in the uplink and downlink data files. For both links, the majority of RSS values were less than -78 dBm. The distributions of RSS

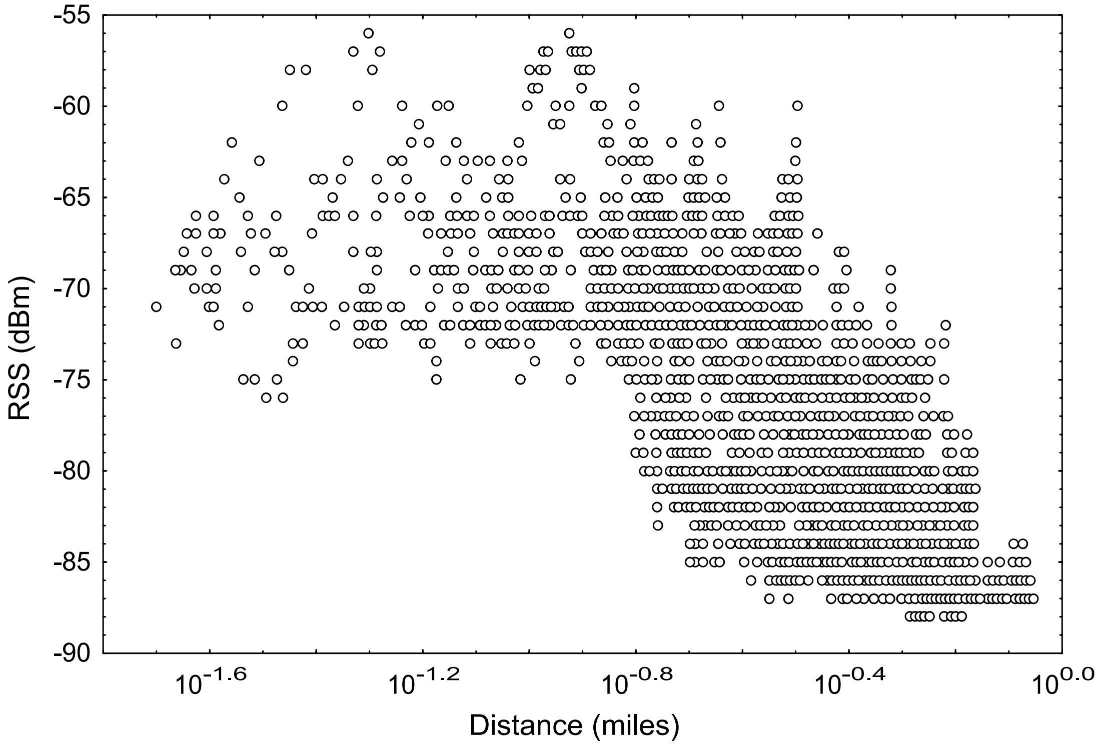

Figure 6.6. Downlink received signal strength (RSS) vs. distance (TAG 7, WCO cell).

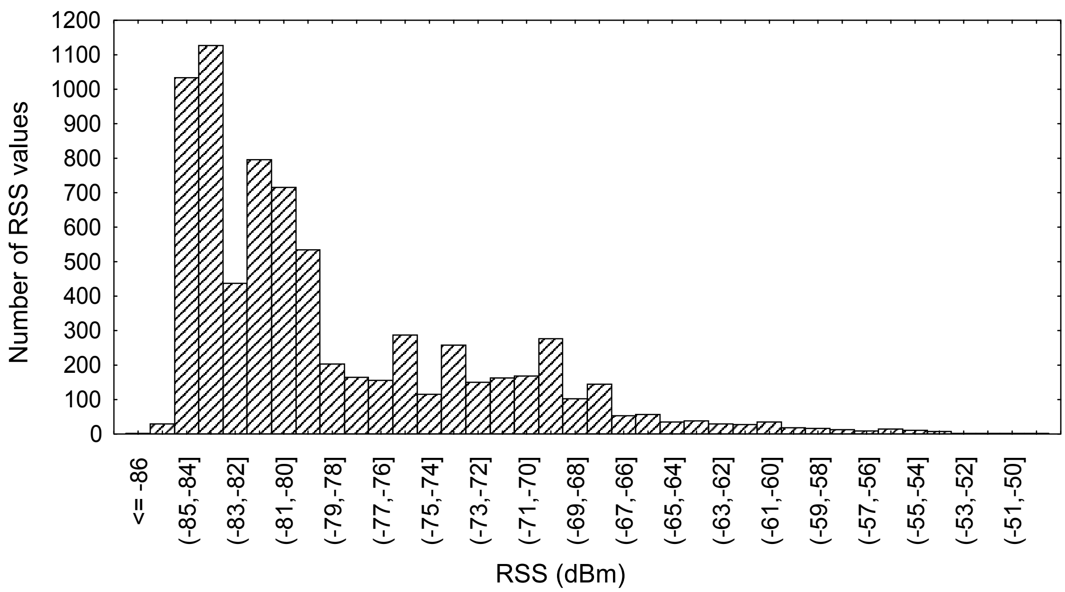

Figure 6.7. Histogram of uplink received signal strength (RSS; TAG 7, WCO cell).

values for the uplink and downlink are similar for RSS values of -78 dBm or greater. The lowest recorded RSS was -88 dBm for the downlink and -85 dBm for the uplink.

Figures 6.9 and 6.10 show the uplink and downlink BER histograms, respectively. The uplink BER histogram for the WCO cell is very similar to the one for the TMCO cell.

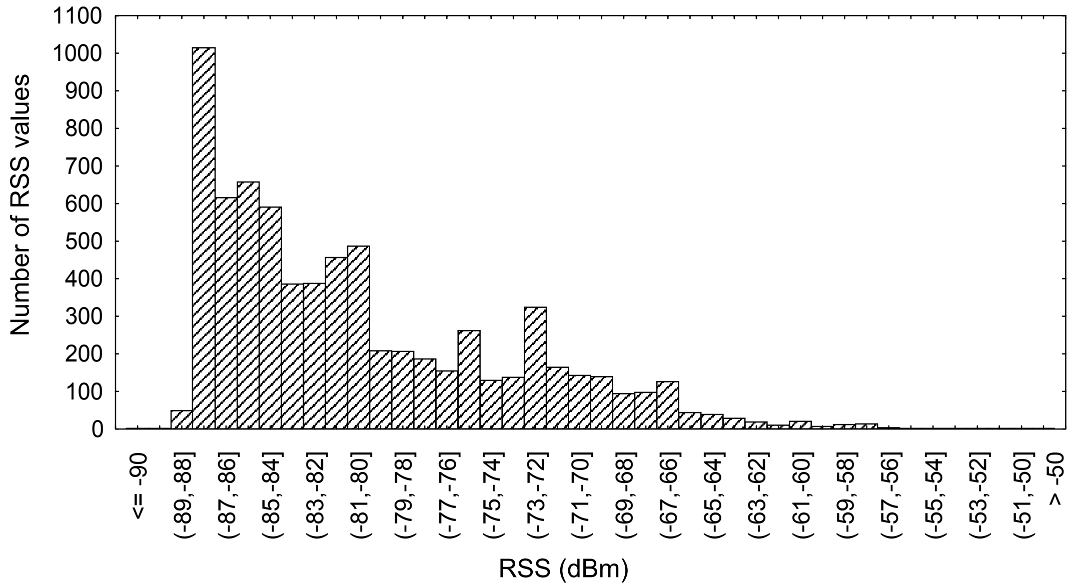

Figure 6.8. Histogram of downlink received signal strength (RSS; TAG 7, WCO cell).

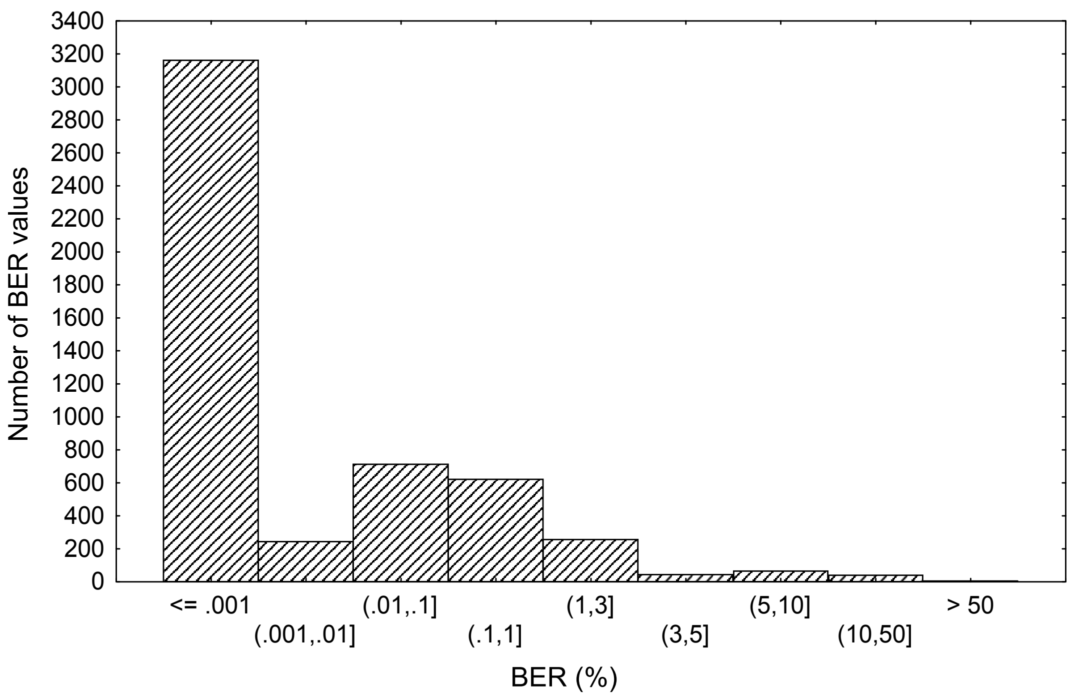

Figure 6.9. Histogram of uplink bit error rate (BER; TAG 7, WCO cell).

Likewise, the downlink BER histogram for the WCO cell is very similar to the one for the TMCO cell. For the uplink, the vast majority of BER values were less than or equal to 0.001%. For the downlink, while a large number of BER values were less than or equal to 0.001%, there were more BER values in the 0.01 - 0.1% and 0.1 - 1.0% ranges than for the uplink.

Figure 6.11 shows a plot of BER as a function of RSS for both the uplink and downlink cases. This plot includes data from both the WCO and TMCO cells. Note that as expected, the BER decreases as the RSS increases. There is a difference between the uplink and downlink curves that is particularly noticeable for BER values of 1% or greater. For a given BER of 1% or greater, the RSS is several dB lower for the uplink case than the downlink case.

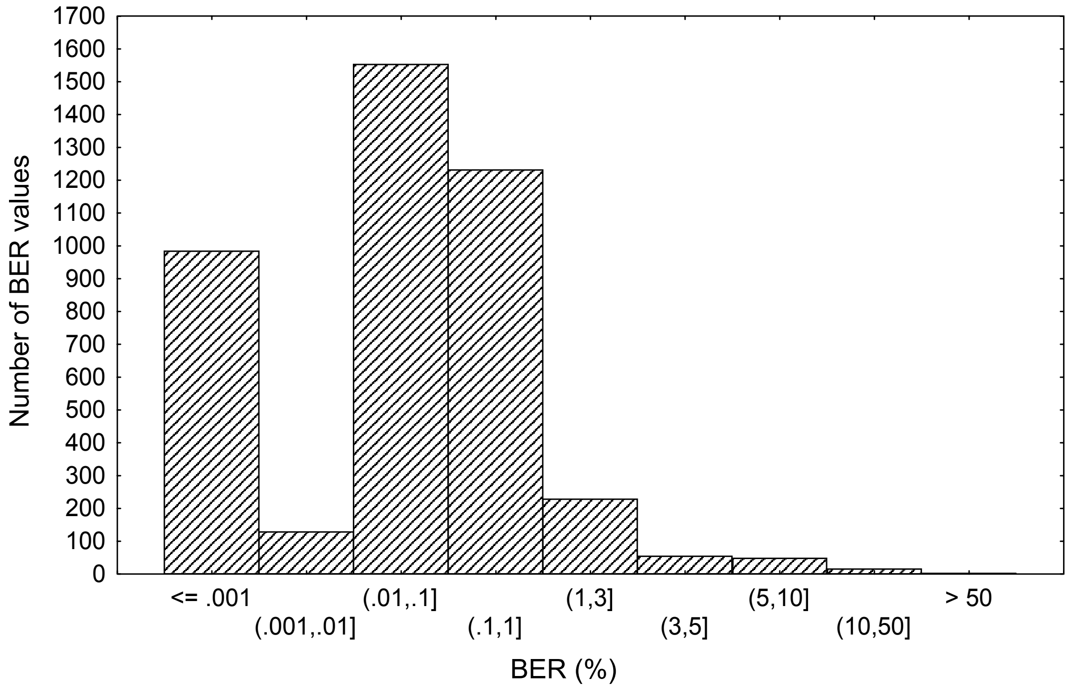

Figure 6.10. Histogram of downlink bit error rate (BER; TAG 7, WCO cell).

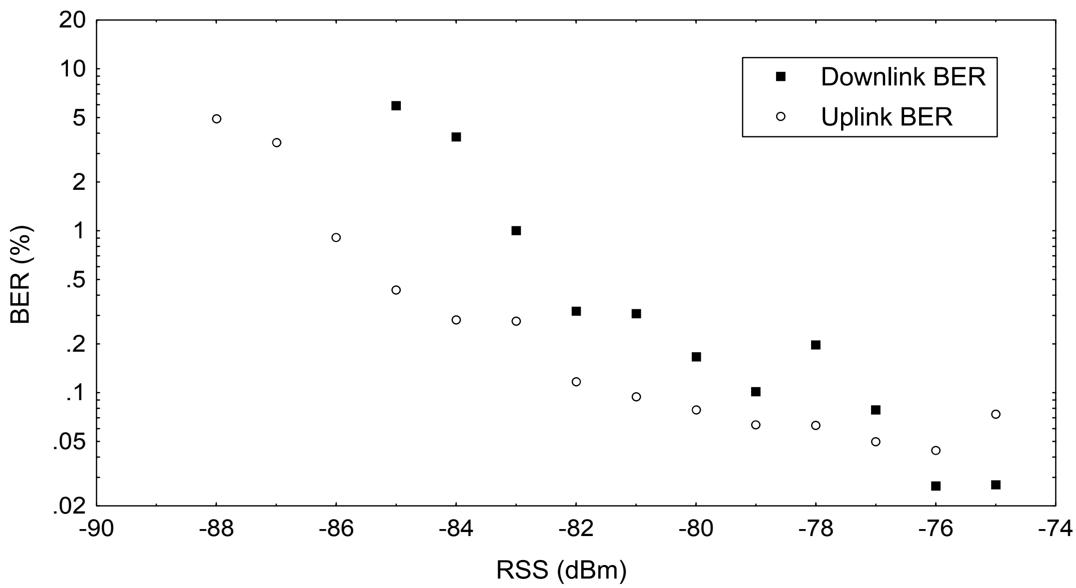

Figure 6.11. Bit error rate (BER) vs. uplink and downlink received signal strength (RSS) for the TMCO and WCO cells (TAG 7).

6.4 Handoff Testing

No handoff testing was performed because only one base station and one mobile unit were used for the TAG 7 testing. This allowed only one sector of one cell to be active at any time during the TAG 7 testing.

6.5 Interference Testing

No interference testing was performed during the TAG 7 testing.

6.6 Voice Quality

As discussed in Section 3.6, two types of voice quality measurements were made for the PCS JTC technology field trials in general: quasi-stationary measurements and handoff measurements. Only the quasi-stationary measurements were performed for the Wideband CDMA (TAG 7) technology; handoff was not possible because only one base station and one mobile unit were used for the TAG 7 testing.

6.6.1 Quasi-stationary Measurements

Voice recordings and various objective measures including uplink and downlink RSS and uplink and downlink BER were collected at locations on a 200-m grid that encompassed the expected coverage area for the TMCO and WCO sites. Note that this grid is smaller than the 0.5-mi grid used in the previous JTC PCS technology field trials. Measurements were taken at the 61 locations shown on the map in Figure 6.12. At each location, data were collected as the measurement van traveled at a given speed. The same vehicle speed and sample time for making the voice recordings was used at each location. The measurement van traveled 10 m over the sample time.

The measurements were taken at each location by establishing a call between the mobile and landline telephones. While the measurement van was in motion, an audio source tape was transmitted over the radio link for both the uplink and downlink simultaneously. The source tape transmitted over each link was the same as that used for TAG 5 testing (see Section 3.6.1).

The received voice transmissions were recorded on digital audio tape simultaneously at the receiver for the uplink and at the receiver for the downlink. The recorded voice segments were then digitized with 16-bit resolution at a 22-Ksample/s rate and stored on a hard disk drive. At each location, after the test run was made to make the voice recordings, another test run was made to collect uplink and downlink RSS and uplink and downlink BER data.

For the quasi-stationary measurements, voice quality of the voice segments was determined by both MOS and expert listener techniques. The following sections discuss these techniques and present the results based on the application of these techniques.

6.6.2 Mean Opinion Score Assessment

To accomplish the MOS testing, a pool of 31 subjects was recruited from the Boulder, Colorado area. Each of the following age groups were represented within the subject pool: 18-25, 25-35, 35-45, 45-55, and those over 55 years of age. There were 17 male and 14 female subjects. The subjects were cordless, noncellular telephone users.

Three groups consisting of eight subjects each and one group consisting of seven subjects were formed from the subject pool. The subjects were asked to rate voice segments by answering the three questions listed in Section 3.6.2 after each segment was presented.

First, the subjects in each of the 4 groups were presented 10 practice voice segments to rate. The practice segments included one 64-kbps wireline voice segment, two voice segments collected from field measurements (one with a definitely acceptable expert listener rating and one with an unacceptable expert listener rating), and seven modulated noise reference unit (MNRU) segments. The MNRU segments were used at the request of the TAG 7 vendor. MNRU's have been used in subjective tests as a reference condition with impairment that sounds similar to the impairment in the coders being tested. The impairment produced by the MNRU segments is defined as the ratio in dB of speech power to speech correlated noise power. For the TAG 7 testing, the MNRU's were produced by artificially adding Gaussian noise of varying levels to the 64-kbps wireline voice segment. The seven MNRU voice segments used during the rating of the practice voice segments included those with a signal-to-noise ratio (SNR) of 5, 10, 15, 20, 25, 30, and 40 dB.

After the practice segments were presented, the subjects in each of the four groups were asked to rate 3 reference voice segments and one quarter of the voice segments taken from the field trial measurements. Every listener within a group listened to the same voice segments. The reference voice segments consisted of one 64-kbps wireline voice segment and two MNRU voice segments with a 15-dB and 25-dB SNR. The voice segments from the field trial measurements came from the uplink or downlink measurements at a portion of the 61 measurement locations.[10] The number and type (male or female) of sentences that were used and the order in which they were used to form a voice segment are the same as that described in Section 4.6.2. Subjects were given two breaks during each session. After all segments were presented, subjects filled out a post-trial questionnaire.

For each voice segment, voice quality ratings (answers to the question "How would you rate the overall quality of the sound?") from each subject within a group were averaged to obtain an MOS. The results from all four of the groups, from a total of 120 voice segments are shown in the histogram in Figure 6.13. Overall, the voice segments were rated favorably, with 70% of the segments rated between fair and excellent. The average MOS was 3.55 and the standard deviation was 1.05.

Figures 6.14 and 6.15 show histograms of MOS's for the uplink and downlink, respectively. The average MOS for the uplink was 3.58 and for the downlink was 3.51. A t-test revealed that there is no statistically significant difference in the average MOS's between the uplink and downlink.

[10] Note that both an uplink and a downlink measurement were made at most, but not all, of the measurement locations.

The relationship between the MOS's and some of the objective measures was initially investigated by generating some scatter plots. Figure 6.16 shows the relationship between the MOS's and average RSS for both the uplink and downlink combined. For RSS values above approximately -80 dBm, some variation in MOS's is seen but most of the MOS's are consistently 3.0 or higher. Note that the MOS's tend to degrade rapidly for values of RSS less than about -80 dBm.

Figures 6.17 and 6.18 show the relationship between the MOS's and the average BER for both the uplink and downlink, respectively. The average BER never exceeded 0.8% on the downlink and 0.45% on the uplink.

Pearson product-moment correlations were performed to determine the correlation between MOS's and average RSS and MOS's and average BER for the data on the uplink and downlink combined. The correlation coefficient between MOS and average RSS was 0.49 and that between MOS and average BER was -0.46. These correlation coefficients between MOS and averaged objective measures suggest that a strong linear relationship between MOS and the objective measures does not exist. The correlation between the average RSS and the average BER was -0.25. A higher correlation between these measures was expected.

While there does not appear to be a strong linear relationship between MOS and the objective measures, there still may be a strong consistently increasing or decreasing relationship between them. The Spearman rank correlation can be used to determine if a consistently increasing or decreasing trend may exist between MOS and the objective measures. Spearman rank correlations were performed to determine the correlation between the ranks of MOS and the ranks of average RSS and between the ranks of MOS and the ranks of average BER. The Spearman rank correlation coefficient between MOS and average RSS was 0.45 and that between MOS and average BER was -0.54. These rank correlations between MOS and averaged objective measures suggest that a strong consistently increasing or decreasing relationship between MOS and the objective measures does not exist.

Note that as in the data analysis of the previous JTC PCS technologies (TAG 5, TAG 2, and TAG 4), the objective measures were averaged over the entire length of the voice segment. By analyzing the instantaneous variation or possibly minimum and maximum values of the objective measures within the voice segment, further insight may be gained on the behavior of MOS's.

By gathering listeners' comments from post-test questionnaires, more information about the nature of MOS's was obtained. Namely, it is evident from questionnaires that there are several types of distortions in quality possible in the TAG 7 voice samples according to listeners:

1) bursts of loud static,
2) "whine" or "feedback noises," and
3) "background hiss."

Figure 6.12. Quasi-stationary measurement locations for TAG 7.

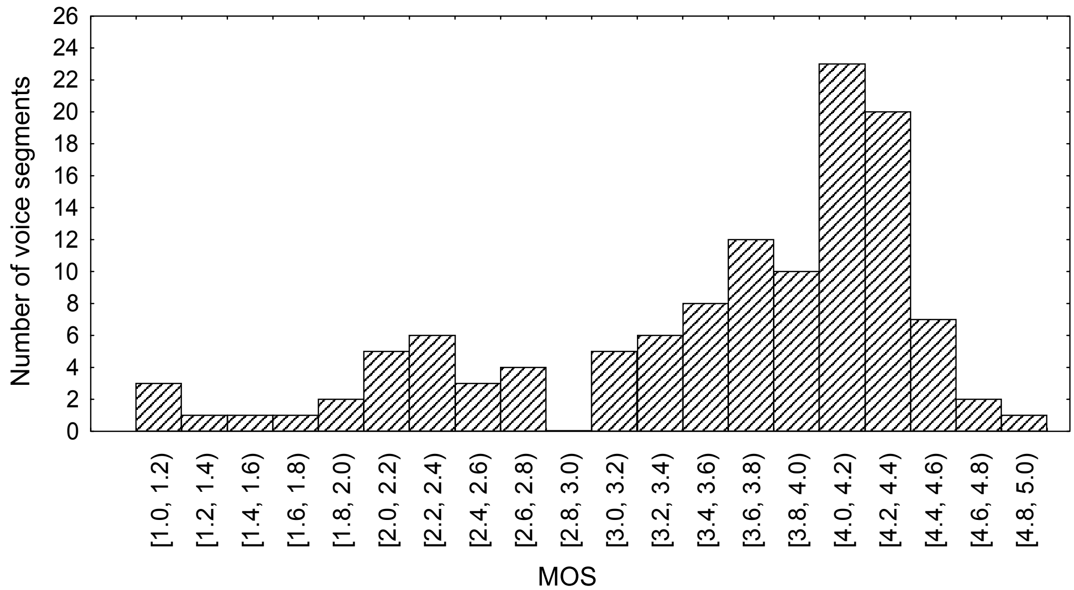

Figure 6.13. Histogram of mean opinion scores (MOS's; TAG 7).

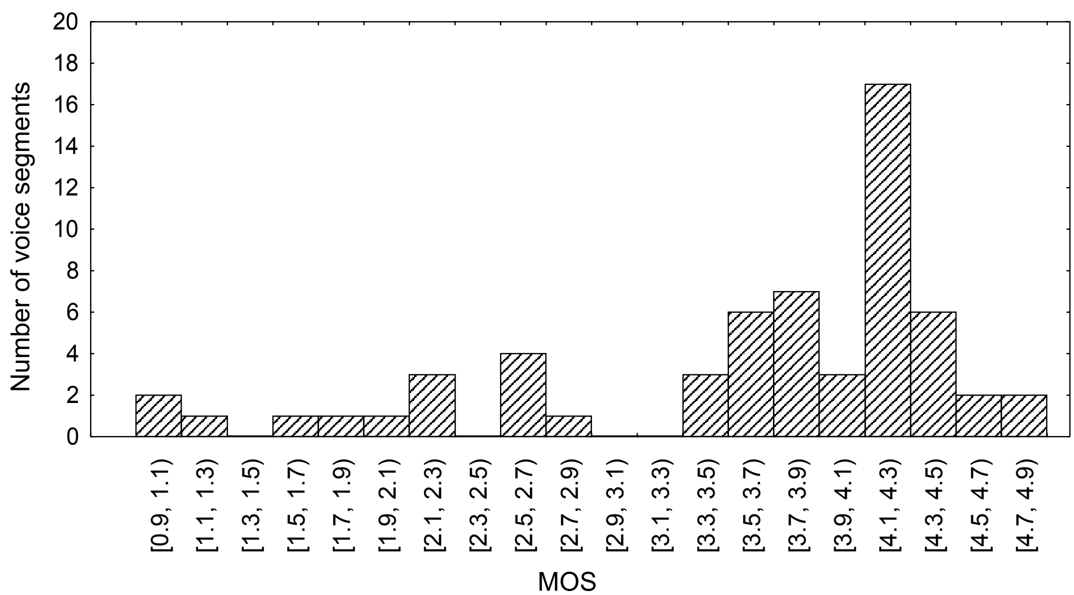

Figure 6.14. Histogram of uplink mean opinion scores (MOS's; TAG 7).

The nature of these distortions are likely judged differently by different listeners. Intelligibility and speaker recognition are two main aspects of perceived quality. For most of the voice samples, intelligibility remained high. As a result, overall MOS's were high.

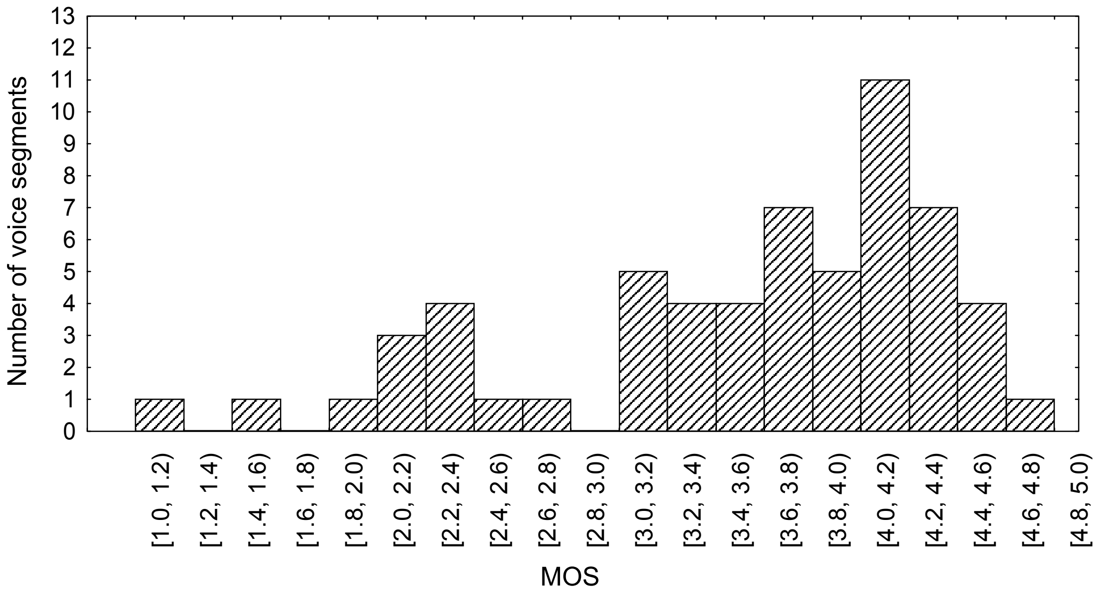

Figure 6.15. Histogram of downlink mean opinion scores (MOS's; TAG 7).

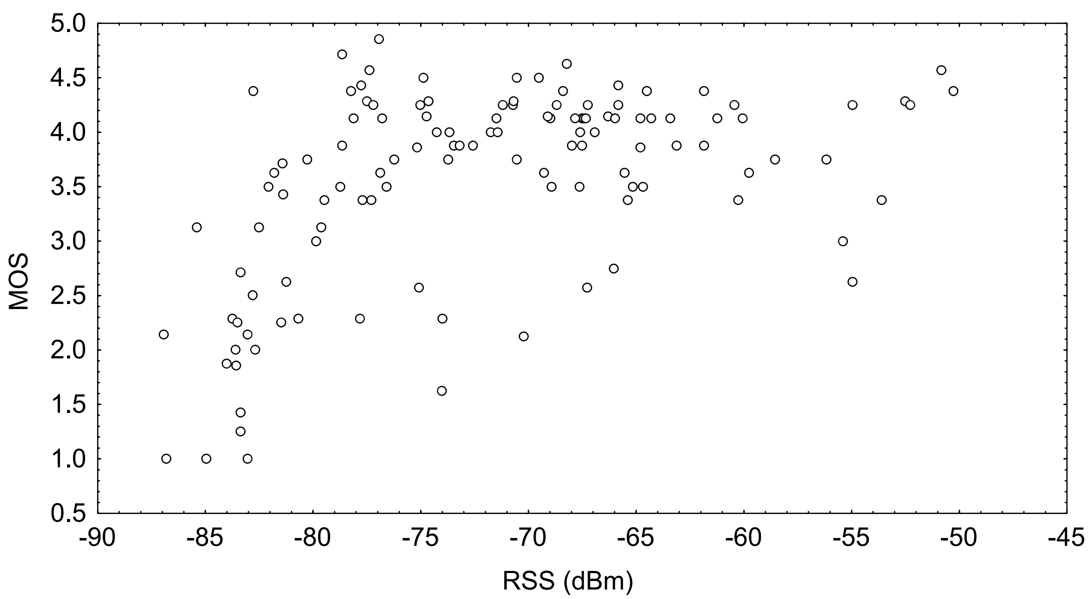

Figure 6.16. Mean opinion score (MOS) vs. average received signal strength (RSS; TAG 7).

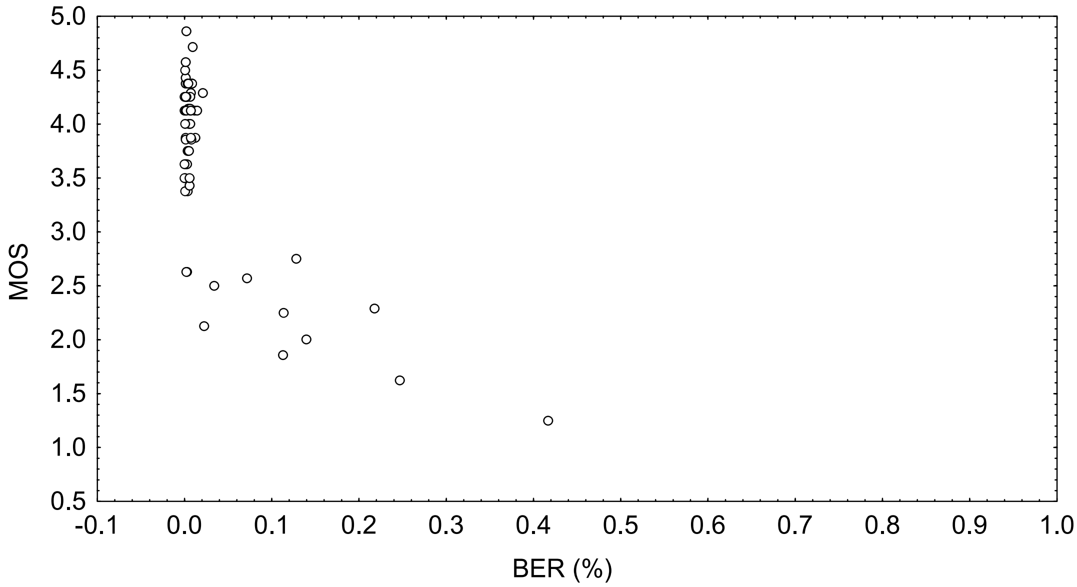

Figure 6.17. Mean opinion score (MOS) vs. average uplink bit error rate (BER; TAG 7).

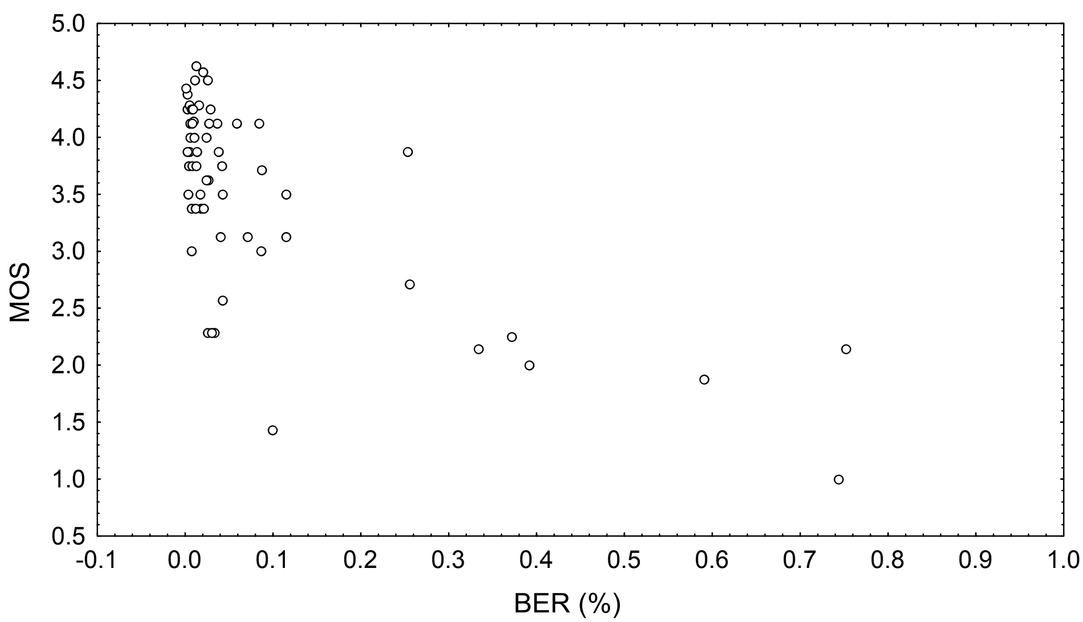

Figure 6.18. Mean opinion score (MOS) vs. average downlink bit error rate (BER; TAG 7).

6.6.3 Expert Listener Assessment

In addition to being rated by listener panels in MOS testing, the voice segments were rated by an expert listener.[11] The expert listener ratings followed the identical procedure as in the PCS 1900 (TAG 5) testing. This procedure is described in Section 3.6.3.

Figure 6.19 shows the relationship between expert listener ratings and percent acceptability (the percentage of listeners rating a given voice segment as acceptable). The boxes represent the middle half of the data (from the 25th percentile to the 75th percentile). The solid circles represent the median percent acceptabilities for each of the expert listener ratings. The lines extending out of the boxes depict the spread of the data.

The expert listener ratings were very good indicators of percent acceptability for the voice segments in the TAG 7 testing. The expert listener ratings accurately predicted the percent acceptability for all of the 120 voice segments recorded.

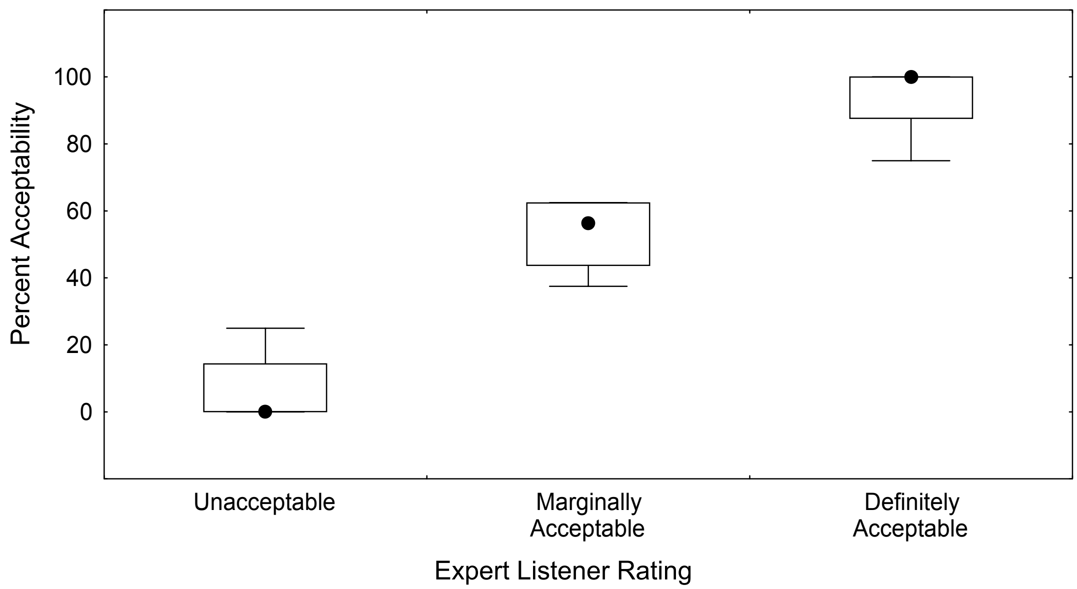

Figure 6.19. Percent acceptability vs. expert listener rating (TAG 7).

The Pearson product-moment correlation coefficient between MOS and percent acceptability was 0.91; this indicates a strong correlation between these measures, as would be expected. The Pearson product-moment correlation coefficient between MOS and expert listener rating was 0.89 and that between percent acceptability and expert listener rating was 0.98; this indicates strong correlation between these measures.

[11] A new expert listener was trained and rated voice segments for the TAG 7 and subsequent JTC PCS technology field trials (TAG 3 and TAG 1). The original expert listener rated voice segments for all of the previous JTC PCS technology field trials (TAG 5, TAG 2, and TAG 4).

6.6.4 Voice Quality Handoff Measurements

Voice quality handoff measurements were not made for the Wideband CDMA (TAG 7) technology field trials.

6.7 Manufacturer's Statement

The statement provided by OKI Electric Industry Co. Ltd. is included in this section. This statement is identical to that given in [5], except for some minor editorial changes.

OKI Electric Industry Co. Ltd. would like to thank U S West, NTIA, and ITS for their support during the demonstration of the performance of the Wideband CDMA radio air interface at the BITB in Colorado. We found the atmosphere and attitude during the tests positive and cooperative.

The purpose of this test was to prove the air-interface functionality and to ascertain coverage area and speech quality. These tasks were successfully accomplished.

Noncommercial hardware was used for the testing. The base station and portable station (mobile unit) transmitter output power was 200 mW, which met the specification. The adaptive pulse code modulation (ADPCM) speech quality was excellent when the portable station (mobile unit) was inside the coverage area. The Wideband CDMA coverage area was smaller than the specification because of the following hardware imperfection:

The RF shielding was imperfect. The resulting 1.9-GHz leakage produced a 2-kHz beat frequency signal which reduced the tracking capability of the RAKE[12] receiver. The degraded receiver sensitivity reduced the coverage radius to less than half of that expected. The Wideband CDMA MOS value was also lower than the specification because of the same hardware imperfection.

[12] The RAKE receiver uses multiple receivers to receive the strongest multipath components of a signal. The receiver then combines these multipath components to provide an improved received signal.